# JMIR Cardio

Electronic, mobile, digital health approaches in cardiology and for cardiovascular health
Volume 9 (2025) ISSN 2561-1011 Editor in Chief: Andrew J Coristine, PhD, Scientific Editor at JMIR Publications, Canada

## Contents

### Original Papers

## Viewpoint



## Review



## Corrigenda and Addenda

## Research Letter

# Self-Reported Acceptance of a Wearable Activity Monitor in Persons With Stroke: Usability Study

Jamie Nam[1], MS, MBA; Grace C Bellinger[1], MPH, PhD; Junyao Li[1,2], BS; Margaret A French[1,3], PT, DPT, PhD; Ryan T Roemmich[1,4], PhD

[1]Department of Physical Medicine and Rehabilitation, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[2]Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[3]Department of Physical Therapy and Athletic Training, University of Utah, Salt Lake City, UT, United States

[4]Center for Movement Studies, Kennedy Krieger Institute, Room 240, 716 N Broadway, Baltimore, MD, United States

**Corresponding Author:**
Ryan T Roemmich, PhD
Department of Physical Medicine and Rehabilitation, Johns Hopkins University School of Medicine, Baltimore, MD, United States

## Abstract

**Background:** Wearable activity monitors offer clinicians and researchers accessible, scalable, and cost-effective tools for continuous remote monitoring of functional status. These technologies can complement traditional clinical outcome measures by providing detailed, minute-by-minute, remotely collected data on a wide array of biometrics, including physical activity and heart rate. There is significant potential for the use of these devices in rehabilitation after stroke if individuals will wear and use the devices; however, the acceptance of these devices by persons with stroke is not well understood.

**Objective:** This study investigated the self-reported acceptance of a commercially available, wrist-worn wearable activity monitor (the Fitbit Inspire 2; Fitbit Inc) for remote monitoring of physical activity and heart rate in persons with stroke. We also assessed relationships between reported acceptance and adherence to wearing the device.

**Methods:** Sixty-five participants with stroke wore a Fitbit Inspire 2 for 3 months, at which point we assessed acceptance using the Technology Acceptance Questionnaire (TAQ), which includes 7 dimensions: perceived usefulness, perceived ease of use, equipment characteristics, privacy concerns, perceived risks, facilitating conditions, and subjective norm. We then performed Spearman correlations to assess relationships between acceptance and adherence to device wear, calculated as both the percentage of daily wear time and the percentage of valid days the device was worn during the 3 weeks preceding TAQ administration.

**Results:** Most participants reported generally agreeable responses, with high overall total TAQ scores across all 7 dimensions, indicating strong acceptance of the device; “Agree” was the median response to 29 of the 31 TAQ statements. Participants generally found the device beneficial for their health, efficient for monitoring, easy to use and to don and doff, and unintrusive to daily life. However, participant responses on the TAQ did not show significant positive correlations with measures of actual device wear time (all $P>.05$).

**Conclusions:** This study demonstrates generally high self-reported acceptance of the Fitbit Inspire 2 among persons with stroke. Participants reported general agreement across all 7 TAQ dimensions, with minimal concerns interpreted as being directly relatable to poststroke motor impairment (eg, donning and doffing the device, using it independently). However, the high self-reported acceptance scores did not correlate positively with measures of real-world device wear. Accordingly, it should not be assumed that persons with stroke will adhere to wearing these devices simply because they report high acceptability.

***(JMIR Cardio 2025;9:e70007)*** doi:10.2196/70007

**KEYWORDS**

stroke; wearables; mobility; technology acceptance; physical activity; remote monitoring; rehabilitation

## Introduction

### Background

Wearable devices have the potential to advance how clinicians and researchers measure functional status by offering accessible, scalable, and cost-effective remote monitoring tools [1,2]. Traditional outcome measures provide only a discrete snapshot of an individual’s functional status [3-5]; emerging wearable technologies have the potential to address this issue by providing minute-by-minute, longitudinal data on physical activity, heart rate, and many other biometrics from daily life that extend beyond clinical or laboratory settings. The use of wearable devices for remote monitoring could offer additional

observational data on functional status measured directly in an individual's daily environment.

Notably, wearable devices could provide valuable insight into recovery following neurologic damage (eg, after stroke) by enabling granular, longitudinal assessment of activity, one of the components of the World Health Organization's International Classification of Functioning, Disability, and Health model [6]. Approximately 80% of persons with stroke experience some form of motor impairment, and around 50% continue to have significant functional limitations 6 months poststroke [7-9]. These limitations often translate into reduced daily activity, as persons with stroke generally walk approximately half as many steps each day as individuals without stroke [10]. Wearable devices provide an opportunity to monitor these functional changes remotely and to generate insights on daily physical activity. However, the ability to perform remote monitoring after stroke using wearable devices is dependent on how well persons with stroke accept these devices.

### Previous Work

Wearable activity monitors such as Fitbit devices generally show high acceptance among healthy individuals [11] and across a range of patient groups, including older adults with cognitive and motor impairment [12,13]. For older adults with cognitive impairment, Fitbit-based interventions may improve motivation for physical activity and sleep, but success is dependent on interfaces that are easy to use [14], reduce cognitive load [15], and are beneficial to the user [14]. In studies of persons with minimal motor impairment, the lack of sustained Fitbit usage and challenges to acceptance have largely been attributed to technical issues—including empty batteries, broken devices, or lost devices [16]—rather than usability concerns. For example, users with multiple sclerosis reported frustration when syncing data between devices [17].

To measure acceptance of wearable technologies across diverse clinical populations, researchers have extensively used the technology acceptance model due to its simplicity and strong empirical support [18,19], with the Technology Acceptance Questionnaire (TAQ) [13] serving as an extension of this framework with additional factors related to user acceptance. Numerous studies have validated this framework across many contexts, demonstrating its predictive power in understanding technology adoption behaviors [18-22]. While wearable activity monitors have demonstrated varying levels of acceptance across the general population, their perceived usability and effectiveness in individuals with stroke remain less explored, which prompted our use of the TAQ to assess their suitability in this specific patient group [23].

### Objective

We aimed to assess the acceptance of wearable devices in persons with stroke. We also examined the relationship between acceptance and adherence to wearing the Fitbit device. We studied Fitbit devices specifically because they have been used extensively for remote monitoring of step count, heart rate, and energy expenditure—among other metrics—in many populations [24]. We hypothesized that (1) acceptance of the Fitbit devices would be variable across persons with stroke but would generally indicate that these technologies are acceptable, useful, and easy to use, and (2) acceptance would be significantly associated with real-world adherence to wearing the Fitbit device.

## Methods

### Recruitment

We recruited persons with stroke from the outpatient rehabilitation clinics at Johns Hopkins Hospital through clinician referrals and MyChart (Epic Systems) messages. The inclusion criteria for this study were: age 18 years and older; history of stroke (confirmed by International Classification of Diseases, Tenth Revision [ICD-10] codes); ownership of a smartphone and in-home Wi-Fi access; walking as a primary form of mobility (with assistive devices allowed); and ability and willingness to install the Fitbit mobile app on a smartphone.

After obtaining consent, the study team asked participants to report their age, sex, race, ethnicity, handedness, and degree of impairment. We then mailed a Fitbit Inspire 2 (Fitbit Inc) to participants and asked them to wear it for 1 year as part of a larger study [25]. This study focuses on a subanalysis of the first 3 months of device use, during which the TAQ was administered to assess participants' experiences and perceptions of the device at the 3-month time point. As an incentive, we allowed participants to keep the Fitbit following completion of the study.

### Study Instructions

We instructed participants to wear the Fitbit on their nonparetic (ie, less impaired) wrist throughout the entire day, including during sleep; if participants had difficulty donning the device and lacked available assistance, we permitted them to wear it on the paretic wrist. We documented the paretic side and the wrist on which the Fitbit was worn. Then, we guided the participants over the phone to set up the device and install the Fitbit app on their smartphone. We instructed participants to remove the device only when showering or charging it. We also asked them to synchronize the device at least once per day using the Fitbit smartphone app. Data from the Fitbit were extracted by a custom-built app, described elsewhere [25,26]. The study team incorporated notifications and reminders to assist with adherence to wearing the Fitbit, as detailed in our previous work [25].

After a participant was enrolled for 3 months, the study team attempted to contact the participant up to 3 times to administer the TAQ, our metric of acceptance. Each contact attempt was documented in Research Electronic Data Capture (REDCap; Vanderbilt University), which included the date and time of the call, the outcome of the attempt (eg, reached, voicemail, and no answer), the participant's response to the assessment, and any notes or follow-up actions required. If the participant was reached, the team member administered the questionnaire verbally. However, if the participant was not reached after 3 attempts, no further attempts were made to administer the TAQ.

### Technology Acceptance Questionnaire

We used the TAQ—which includes dimensions for perceived usefulness (PU), perceived ease of use (PEOU), equipment characteristics, privacy concerns, perceived risks, facilitating conditions, and subjective norm—as established by Puri et al [13]. The full questionnaire is described in Puri et al [13]. The TAQ consisted of 31 statements rated on a 5-point Likert scale where participants indicated their levels of agreement or disagreement with each statement. These 31 statements covered dimensions of PU (n=5 statements), PEOU (n=7 statements), facilitating conditions (n=2 statements), subjective norm (n=3 statements), equipment characteristics (n=8 statements), privacy concerns (n=3 statements), and perceived risks (n=3 statements). We summed the scores within each dimension ("Strongly disagree"=1 point, "Disagree"=2 points, "Neutral"=3 points, "Agree"=4 points, "Strongly agree"=5 points). We note that, to ensure the reliability and validity of the measure, certain statements in the TAQ were negatively framed to mitigate acquiescence bias [27]. As is standard, we minimized possible response biases in which higher numerical values represent lower agreement by applying a reverse-coding procedure to these statements: each affected score was standardized by subtracting it from 6 for statements 2, 10, 17, 25, and 27. This was done before summing the scores within each dimension. We then calculated the total TAQ score by summing the scores from all items across all dimensions. In addition to the 31 statements, there are also 6 multiple-choice questions addressing various aspects of device use that are not assigned to any dimension. Finally, we also provided participants with the opportunity to share open-ended comments about their experiences using the devices.

### Fitbit Adherence

To assess adherence, we analyzed Fitbit data from the 21 days immediately preceding completion of the TAQ. This timeframe was selected because the TAQ items specifically reference user perceptions of their Fitbit over the previous 3 weeks. We identified Fitbit wear minutes using the accelerometry package [28] in R (R Foundation for Statistical Computing) [29]. We calculated 2 primary adherence metrics. First, we calculated the average percentage of time the device was worn each day by dividing the number of wear minutes by the total minutes in a day (1440), multiplying by 100, and averaging the daily percentages across the 21-day period. Second, we calculated the percentage of valid wear days as the number of days within the 3-week window with at least 75% of 24-hour wear time (ie, 1080 or more minutes). The number of valid days was then divided by 21 (the total days in the assessment window) and multiplied by 100 to determine the percentage of valid days.

### Statistical Analysis

We calculated descriptive statistics for the 37 individual items of the TAQ (inclusive of the 31 Likert scale statements and 6 additional multiple-choice questions), the summed scores on each of the 7 dimensions from the TAQ, and the summed total score of the TAQ. We also report descriptive statistics for the metrics of adherence. To assess relationships between Fitbit adherence and TAQ responses, we used Spearman correlations (due to the ordinal nature of the TAQ responses). We performed all statistical analyses using R (version 4.4.1; R Foundation for Statistical Computing) [29] with $\alpha$=.05.

### Ethical Considerations

All participants provided oral consent, as approved by the Institutional Review Board at the Johns Hopkins University School of Medicine (IRB00247292). All data were deidentified. We did not provide participants with monetary compensation; however, they were allowed to keep their Fitbit devices following completion of the study as compensation for their participation.

## Results

### Participants

Of the 108 persons with stroke contacted, 98 participants enrolled in this study (8 individuals either did not meet the inclusion criteria or declined to participate and 2 individuals enrolled in the larger study but declined the Fitbit portion). Among the enrolled participants, 4 were lost to follow-up (ie, they could not be reached despite repeated attempts during the 3-month study period) and 8 withdrew voluntarily (ie, they chose to discontinue their involvement after enrolling). Furthermore, 21 participants did not complete the TAQ 3 months postenrollment. As a result, we included 65 participants with stroke in the final analysis. We present a participant enrollment flow diagram in Figure 1 and summarize the characteristics of participants included and excluded in the analysis in Table 1.

**Figure 1.** Flow diagram showing participant enrollment and participation at various stages of the study. TAQ: Technology Acceptance Questionnaire.

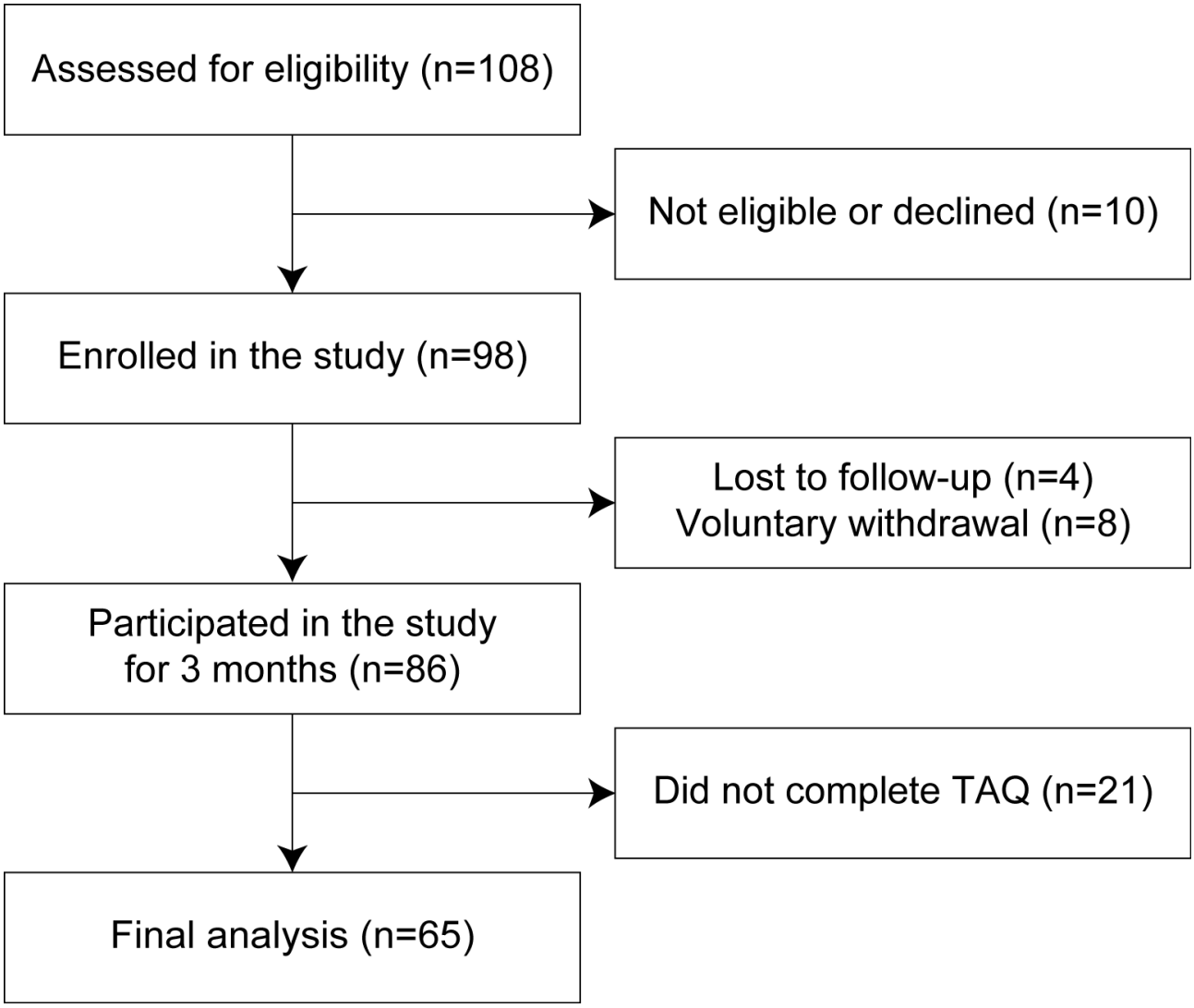


**Table .** Study participant characteristics.

| Characteristic | Included participants (n=65) | Excluded participants (n=35)[a] |
|---|---|---|
| Age (years), mean (SD) | 62.4 (12.4) | 60.5 (12.5) |
| Days poststroke at date of enrollment, mean (SD) | 1053.9 (1664.1) | 230.1 (367.0) |
| Days between the date of enrollment and the date of TAQ completion, mean (SD) | 127.8 (42.9) | —[b] |
| Sex, n (%) | | |
| Male | 37 (56.9) | 20 (57.1) |
| Female | 28 (43.1) | 15 (42.9) |
| Race, n (%) | | |
| American Indian or Alaska Native | 2 (3.1) | 0 (0.0) |
| Asian | 3 (4.6) | 1 (2.9) |
| Black or African American | 21 (32.3) | 22 (62.9) |
| White or Caucasian | 38 (58.5) | 10 (28.6) |
| Multiple | 1 (1.5) | 2 (5.7) |
| Ethnicity, n (%) | | |
| Hispanic or Latino | 4 (6.2) | 1 (2.9) |
| Not Hispanic or Latino | 61 (93.8) | 34 (97.1) |
| Use of an assistive device for walking, n (%) | | |
| Yes | 28 (43.1) | 13 (37.1) |
| No | 37 (56.9) | 21 (60.0) |
| No response | 0 (0.0) | 1 (2.9) |
| Able to move the paretic arm, n (%) | | |
| Yes | 57 (87.7) | 30 (85.7) |
| No | 8 (12.3) | 3 (8.6) |
| No response | 0 (0.0) | 2 (5.7) |
| Able to bring a hand from the lap to the table, then to top of chest, and reach for object above the table surface, n (%) | | |
| Yes | 56 (86.2) | 31 (88.6) |
| No | 8 (12.3) | 2 (5.7) |
| Unsure | 1 (1.5) | 0 (0.0) |
| No response | 0 (0.0) | 2 (5.7) |
| Severe shoulder pain that limits the ability to move the paretic arm, n (%) | | |
| Yes | 12 (18.5) | 3 (8.6) |
| No | 53 (81.5) | 31 (88.6) |
| No response | 0 (0.0) | 1 (2.9) |
| Fitbit worn on paretic or nonparetic wrist[c], n (%) | | |
| Paretic | 15 (23.1) | — |
| Nonparetic | 49 (75.4) | — |
| Unknown | 1 (1.5) | — |
| Fitbit worn on poststroke dominant or nondominant wrist[c], n (%) | | |
| Dominant | 36 (55.4) | — |

| Characteristic | Included participants (n=65) | Excluded participants (n=35)[a] |
|---|---|---|
| Nondominant | 28 (43.1) | — |
| Unknown | 1 (1.5) | — |

[a]The 35 excluded participants include the 2 individuals that enrolled in the larger study but declined the Fitbit portion.

[b]Not applicable.

[c]Wrist placement not available for one included participant.

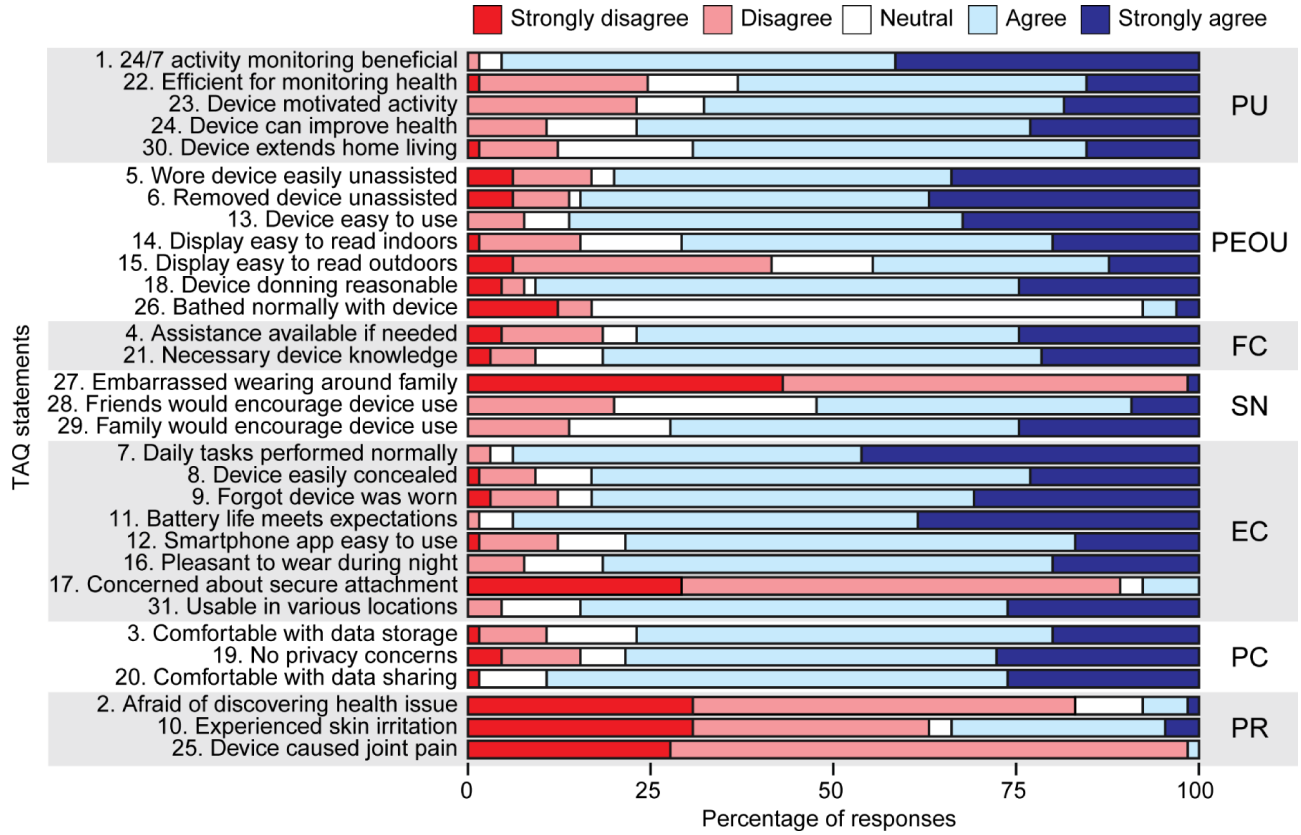
## Self-Reported Acceptance of Wearable Devices in Participants With Stroke

We report percentages of responses (Strongly disagree, Disagree, Neutral, Agree, and Strongly agree) to each of the 31 TAQ statements rated on Likert scales, organized by dimension (Figure 2). The 5 statements with the highest proportions of "Agree" or "Strongly agree" responses were (in order; 2a, 2b, and so forth indicate multiple statements with the same proportions of responses): (1) Statement 1: "I think that monitoring my activity and health 24 hours a day, 7 days a week, can be a good thing;" (2a) Statement 7: "I was able to perform my daily tasks as usual while wearing the device;" (2b) Statement 11: "The battery life of the device meets my expectations;" (4) Statement 18: "I was able to put the device on in a reasonable amount of time;" and (5) Statement 20: "I am comfortable with my health data being shared with equipment manufacturers as long as it is shared anonymously."

The five statements with the lowest proportions of "Agree" or "Strongly agree" responses were (in order): (1a) Statement 25: "Wearing the device caused me to have joint pain," (1b) Statement 27: "I was embarrassed to wear the device in front of family members," (3a) Statement 2: "I was afraid that the device would discover a major health issue," (3b) Statement 17: "I was concerned that the device is not securely attached to me," and (3c) Statement 26: "I was able to shower or bathe normally while wearing the device."

As statements 1a-3b are reverse-coded; thus, a lower proportion of "Agree" or "Strongly agree" responses to these statements indicates a more positive perception of the device. With the exceptions of the 5 statements listed above, as well as statements 10 ("I experience skin irritations while wearing the device") and 15 ("I find the display of the device easy to read outdoors"), a majority of participants responded either "Agree" or "Strongly agree" to the remaining 24 statements. We also note that responses to statement 26 were likely influenced by our instructions to participants to remove the device while bathing or showering.

**Figure 2.** Percentage of responses to each statement on the TAQ grouped by the 7 TAQ dimensions. EC: equipment characteristics; FC: facilitating conditions; PC: privacy concerns; PEOU: perceived ease of use; PR: perceived risks; PU: perceived usefulness; SN: subjective norm; TAQ: Technology Acceptance Questionnaire.



We present group means, standard deviations (SD), medians, and interquartile ranges (IQR) of scores for each of the 7 dimensions in Table 2 (alongside the minimum and maximum possible scores for each dimension) and for each individual TAQ statement in Table 3. Medians of the summed scores for each dimension ranged from 71% in PEOU (median=25, maximum possible score=35) to 80% in all other dimensions (Table 2). Furthermore, all dimensions—with the exception of PEOU—showed median scores of 4 (“Agree”) on each individual statement (Table 3). Modestly lower scores in the PEOU dimension were largely driven by generally less agreeable responses to statement 15 (“I found the display of the device easy to read outdoors”) and statement 26 (“I was able to shower or bathe normally while wearing the device”), again noting that participants were instructed to remove the device before bathing or showering. The median for the total TAQ score was 76% of the maximum possible score (median=118, maximum possible score=155).

**Table .** Statistics of the responses to the Technology Acceptance Questionnaire (TAQ) dimensions and the full TAQ.

| Dimension | Scores, mean (SD) | Scores, median (IQR) | Scores, possible range (min-max) |
|---|---|---|---|
| Perceived usefulness | 19.1 (3.3) | 20 (17-21) | 5-25 |
| Perceived ease of use | 25.7 (3.9) | 25 (23-28) | 7-35 |
| Facilitating conditions | 7.7 (1.5) | 8 (7-8) | 2-10 |
| Subjective norm[a] | 11.6 (1.9) | 12 (10-13) | 3-15 |
| Equipment characteristics[a] | 32.5 (3.5) | 32 (30-35) | 8-40 |
| Privacy concerns | 11.8 (2.0) | 12 (10-13) | 3-15 |
| Perceived risks[a] | 11.9 (2.0) | 12 (10-13) | 3-15 |
| Technology Acceptance Questionnaire | 120.4 (11.8) | 118 (114-128) | 31-155 |

[a]These dimensions have reverse items in the subscore. Perceived risks consists of all reversed questions; therefore, the entire dimension is reversed.

**Table .** Statistics of the responses to each individual statement of the Technology Acceptance Questionnaire (TAQ).

| Statement | Scores, mean (SD) | Scores, median (IQR) |
|---|---|---|
| 24/7 activity monitoring beneficial | 4.4 (0.6) | 4 (4-5) |
| Afraid of discovering health issue[a] | 4.0 (0.9) | 4 (4-5) |
| Comfortable with data storage | 3.8 (0.9) | 4 (4-4) |
| Assistance available if needed | 3.8 (1.1) | 4 (4-4) |
| Wore device easily unassisted | 3.9 (1.2) | 4 (4-5) |
| Removed device unassisted | 4.0 (1.1) | 4 (4-5) |
| Daily tasks performed normally | 4.4 (0.7) | 4 (4-5) |
| Device easily concealed | 4.0 (0.9) | 4 (4-4) |
| Forgot device was worn | 4.0 (1.0) | 4 (4-5) |
| Experienced skin irritation[a] | 3.6 (1.3) | 4 (2-5) |
| Battery life meets expectations | 4.3 (0.6) | 4 (4-5) |
| Smartphone app easy to use | 3.8 (0.9) | 4 (4-4) |
| Device easy to use | 4.1 (0.8) | 4 (4-5) |
| Display easy to read indoors | 3.7 (1.0) | 4 (3-4) |
| Display easy to read outdoors | 3.1 (1.2) | 3 (2-4) |
| Pleasant to wear during night | 3.9 (0.8) | 4 (4-4) |
| Concerned about secure attachment[a] | 4.1 (0.8) | 4 (4-5) |
| Device donning reasonable | 4.0 (0.9) | 4 (4-4) |
| No privacy concerns | 3.9 (1.1) | 4 (4-5) |
| Comfortable with data sharing | 4.1 (0.7) | 4 (4-5) |
| Necessary device knowledge | 3.9 (0.9) | 4 (4-4) |
| Efficient for monitoring health | 3.5 (1.1) | 4 (3-4) |
| Device motivated activity | 3.6 (1.0) | 4 (3-4) |
| Device can improve health | 3.9 (0.9) | 4 (4-4) |
| Device caused joint pain[a] | 4.2 (0.5) | 4 (4-5) |
| Bathed normally with device | 2.8 (0.8) | 3 (3-3) |
| Embarrassed wearing around family[a] | 4.4 (0.7) | 4 (4-5) |
| Friends would encourage device use | 3.4 (0.9) | 4 (3-4) |
| Family would encourage device use | 3.8 (1.0) | 4 (3-4) |
| Device extends home living | 3.7 (0.9) | 4 (3-4) |
| Usable in various locations | 4.1 (0.8) | 4 (4-5) |

[a]These statements have been reverse coded to preserve directionality.

Next, we analyzed responses to the 6 multiple-choice questions from the TAQ that did not have designated dimensions (Table 4). Most participants found the device useful, with 95.4% (62/65) rating the information provided as either “very useful” or “somewhat useful.” Nearly all participants (92.3%, 60/65) expressed willingness to continue using the device to monitor their health, and 96.9% (63/65) reported wearing the device for 15 - 21 days out of the 21-day period. In terms of value, most participants indicated a willingness to pay no more than $100 for the device. Overall, 90.8% (59/65) of participants reported looking at their health data provided by their device. Finally, self-perception of activity levels varied, with 64.6% (42/65) of participants considering themselves to be active.

**Table .** Responses to the Technology Acceptance Questionnaire (TAQ) multiple-choice questions.

| Questions and response options | Respondents, n (%) |
|---|---|
| How useful did you find the information provided by the smart wearable device (such as step count, sleep data, and heart rate) either on the wearable itself or in the smartphone application? | |
| Very useful | 36 (55.4) |
| Somewhat useful | 26 (40.0) |
| Not very useful | 3 (4.6) |
| Not at all useful | 0 (0.0) |
| Would you use the device you used during the last 21 days to continue to monitor or track your physical activity or health? | |
| Yes | 60 (92.3) |
| No | 5 (7.7) |
| Over the last 21 days, how often do you think you wore the smart wearable device? | |
| Never | 0 (0.0) |
| Between 0 and 7 days | 0 (0.0) |
| Between 8 and 14 days | 2 (3.1) |
| Between 15 and 21 days | 63 (96.9) |
| How much would you be willing to pay for the device you wore during the last 21 days? | |
| $0 | 15 (23.1) |
| $1-$50 | 17 (26.2) |
| $51-$100 | 24 (36.9) |
| $101-$200 | 9 (13.8) |
| $201-$300 | 0 (0.0) |
| $300-$400 | 0 (0.0) |
| Did you find yourself looking at your health data in the smartphone application more or less often after the first few days? | |
| No, I looked at the health data fairly consistently throughout the 21-day period | 21 (32.3) |
| Yes, I looked at the health data more often after the first few days of use | 28 (43.1) |
| Yes, I looked at the health data less often after the first few days of use | 10 (15.4) |
| I did not look at my health or am not interested in monitoring it | 6 (9.2) |
| Do you consider yourself to be an active person? | |
| Yes | 42 (64.6) |
| No | 23 (35.4) |

## Relationships Between Self-Reported Acceptance and Fitbit Adherence

Overall, participants wore the Fitbit for an average of 80.0% (SD 24.7%) of the total minutes in a day, with a median wear time of 91% and an IQR of 22%. Wear time exceeded the threshold needed to be considered a valid wear day on 78.0% (SD 25.8%) of days, with a median of 90% and an IQR of 33%. The scatterplots in Figures 3 and 4 show relationships between the summed scores of the different TAQ dimensions (as well as TAQ total scores) and the percentages of wear time and valid wear days (days with ≥1080 wear minutes), respectively. Contrary to our hypothesis, there were no statistically significant positive relationships between TAQ dimension summed scores or total score and Fitbit adherence metrics. We did, however, observe 2 small but statistically significant negative associations: between PU and percent wear time ($\rho=-.27$; $P=.03$) and between PU and the percentage of valid wear days ($\rho=-.26$; $P=.04$).

**Figure 3.** Scatterplots showing relationships between summed scores for each TAQ dimension (as well as total TAQ scores) and Fitbit wear time. Asterisks (*) indicate statistically significant relationships (*P*<.05). EC: equipment characteristics; FC: facilitating conditions; PC: privacy concerns; PEOU: perceived ease of use; PR: perceived risks; PU: perceived usefulness; SN: subjective norm; TAQ: Technology Acceptance Questionnaire.

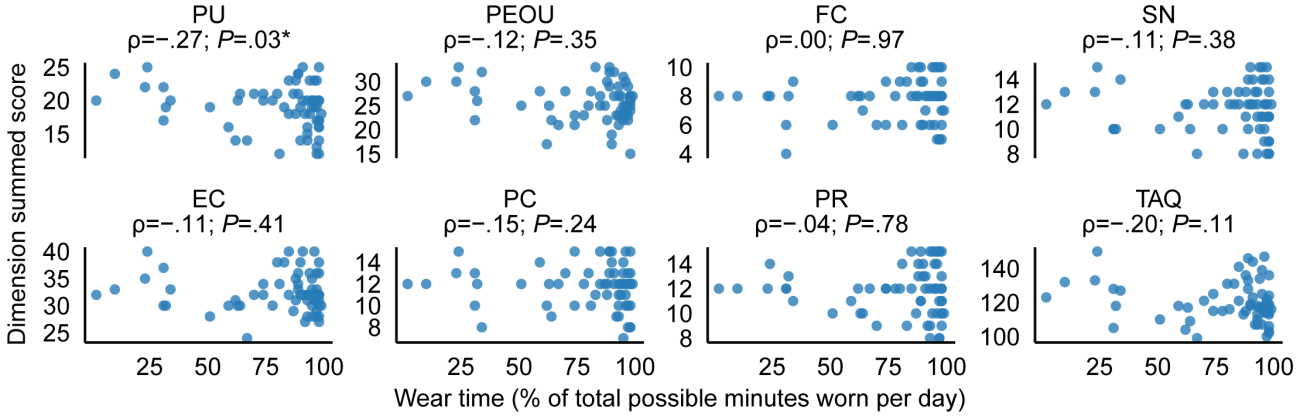
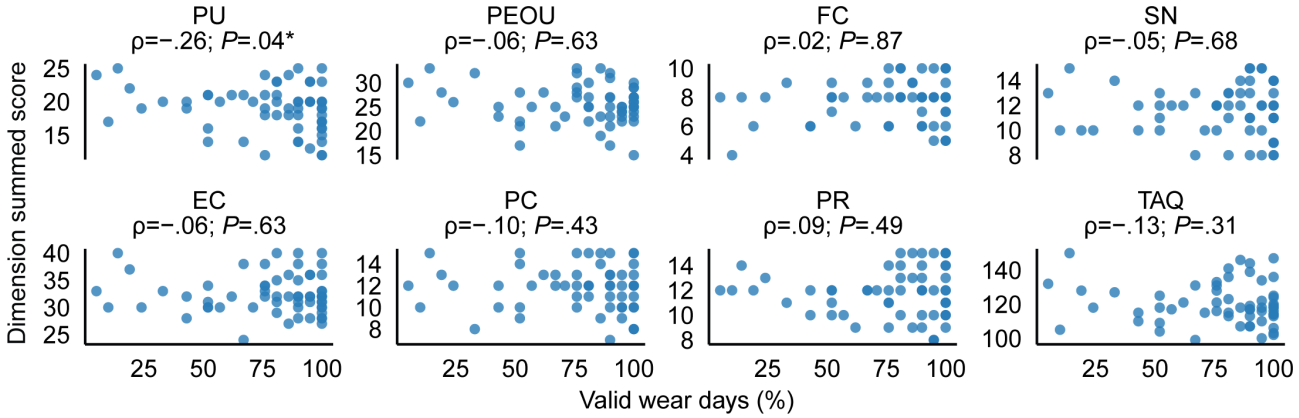



**Figure 4.** Scatterplots showing relationships between summed scores for each TAQ dimension (as well as total TAQ scores) and valid Fitbit wear days. Asterisks (*) indicate statistically significant relationships (*P*<.05). EC: equipment characteristics; FC: facilitating conditions; PC: privacy concerns; PEOU: perceived ease of use; PR: perceived risks; PU: perceived usefulness; SN: subjective norm; TAQ: Technology Acceptance Questionnaire.




### Open-Ended Comments on the Devices

Participants provided a series of positive and negative comments about their experiences with the devices. Themes of positive comments included the ability to monitor real-time heart rate, the use of step count data as motivation to increase activity, and the provision of sleep data. Themes of negative comments included difficulties donning and doffing the device, technical challenges with the associated mobile app, difficulty reading the screen on the device, and concerns that other wearables and smartwatches might provide more relevant or more comprehensive data.

Qualitatively, we did not find any consistent themes in the open-ended comments that related to adherence to wearing the device. For example, among participants who wore the device for fewer than 50% of possible minutes (n=8), we observed that most of the comments were largely positive in nature. These participants reported that the Fitbit provided useful step count information, helped them stay on track with physical activity, and offered useful sleep data. The only negative comments were centered around the potential utility of additional training for using the device and the small size of the device screen.

## Discussion

### Principal Findings

Our study examined self-reported perceptions of the Fitbit Inspire 2 wearable activity monitor among individuals with stroke, as measured by the TAQ. A majority of the participants thought the device was beneficial for their health, efficient for monitoring their health, easy to use and to don and doff, and unintrusive to daily life; one notable exception was the response to the statement “I find the display of the device easy to read outdoors.” Generally, participants did not express significant concerns about privacy or data security, consistent with previous studies [30,31]. Contrary to our hypothesis, more agreeable responses to the TAQ statements were not associated with higher average daily wear time and valid wear days at a statistically significant level.

### Comparison With Previous Work

The findings of our study are consistent with previous literature demonstrating acceptance of wearable activity monitors in other populations [13,32-36]. Given the growing interest in using wearables for activity monitoring and telerehabilitation after stroke [26,37-40], it is important to consider not only their potential benefits but also the potential challenges related to low levels of physical activity and cognitive and motor impairments that may affect this population’s acceptance and

engagement [10,41,42]. There were no commonly reported acceptance concerns that we deemed likely to be related to poststroke motor impairment. For example, participants widely agreed with the statements "I was able to wear the device easily without help from another person," "I find the device easy to use," "I was able to put the device on in a reasonable amount of time," and "I was able to remove the device easily without help from another person." These findings complement recent studies demonstrating the perceived value and user satisfaction of wearable technologies in stroke rehabilitation [43-45] and support a path toward scalable implementation of remote monitoring with wearable devices. This is likely due in part to the flexibility we allowed in permitting participants to wear the device on their paretic side if needed, accommodating individual motor abilities.

We also highlight that the study participants reported generally agreeable responses across all 7 dimensions of the TAQ. Previous work highlighted that technical and usability issues (eg, requiring a mobile app to sync the data from the device to the server) may affect the PU of wearable devices [16]; however, we did not observe this in our sample. This is potentially because any technical difficulties were often addressed via interactions with the study team. Our findings across the different dimensions were largely similar to those reported in a previous sample of older adults [13]. It is important to emphasize that monitoring of device adherence may be necessary despite the high reported acceptance. Our findings did not support the hypothesis that higher user acceptance as measured by the TAQ would correlate with adherence to wearing the device, as we did not observe statistically significant positive correlations between TAQ responses and our measures of adherence. Furthermore, we found that 5 participants reported that they had worn the device for at least 15 of the preceding 21 days (in response to the multiple-choice question) but provided fewer than 15 days of Fitbit data. This revealed that high reported acceptance of the device does not guarantee that a patient or research participant will necessarily adhere to wearing the device in everyday life. Technologies that help to automate oversight of device wear and messaging to promote adherence will be important for ensuring data quality [46].

The correlational analyses indicated small but significant negative associations between the PU score and both adherence metrics. These results are contrary to our hypotheses, which were grounded in the technology acceptance model and related literature, where PU is typically positively associated with adoption behaviors [47,48]. Existing studies have shown that higher PU is often associated with sustained use of technology across various domains [49-51], including health care settings [52-55]. Our findings may be attributed to the specific context in which the wearable device was used. Unlike most previous technology adoption studies where participants voluntarily adopted technology based on its usefulness, our study cohort used the Fitbit as part of their participation in a research study. This mandated context could have influenced PU differently from typical motivational factors driving technology adoption, as individuals may have rationalized their behaviors based on the rewards of participating or the consequences of noncompliance. It is possible that participants did not view the Fitbit as inherently useful for their health recovery goals but instead perceived it as a tool for fulfilling study requirements. They may have also overreported PU due to social desirability (ie, aiming to please the study team). While survey responses suggest that most participants agreed that the device could improve health and monitor well-being efficiently, these endorsements may reflect a general perception of health technology utility rather than a personalized alignment with stroke recovery needs. Consequently, their assessment of PU may reflect this externally driven motivation rather than genuine alignment with personal health management goals.

As the push toward clinical use of wearables in stroke rehabilitation continues to move forward, it is also important to consider the needs and perspectives of all key stakeholders—patients, their family members and caregivers, and clinicians—in addition to the device acceptance reported by participants in our study. Recent studies have provided vital information regarding how persons with stroke prefer to receive data from wearables and identified a set of metrics deemed most useful [23]; incorporated perspectives from patients and clinicians on the value of using wearables and identified preferences for incorporation into clinical care [44]; and demonstrated important design considerations for adoption of the wearables and accompanying smartphone apps as outlined by persons with stroke [45]. For example, considering the difficulty of donning and doffing, an alternative strap such as a Velcro strap instead of the original buckle band may improve usability. Future work should consider these multifaceted aspects—patient acceptance, patient and clinician data provision preferences, and device design—as we move closer to clinical implementation of wearables in stroke rehabilitation.

## Limitations

We acknowledge some limitations in our study. First, our sample was heterogeneous regarding stroke chronicity. Accordingly, we did not design this study to assess how wearable device acceptance may differ across stages of stroke recovery (eg, acute, subacute, and chronic). However, this diversity in stroke recovery stages could be beneficial, as it reflects the clinical reality in which wearable devices in stroke rehabilitation should not discriminate based on recovery stage but rather be accessible and valuable to patients at various points in their recovery. Second, we only used the Fitbit Inspire 2 device. While we anticipate that many of our findings may generalize across different models of commercially available wearable devices due to the nature of the statements included in the TAQ, we do not have data to support this directly. Third, we focused on the TAQ to provide information about device acceptance in particular. We did not use other instruments that could provide additional data on other aspects of patient perceptions about wearable devices (eg, the System Usability Scale for assessment of usability). Finally, our study focuses on individuals who own smartphones and have home Wi-Fi. This digital access criterion may have biased the sample toward more tech-savvy or higher-functioning individuals.

## Conclusion

This study reported the perceived acceptance of a wrist-worn activity monitor among persons with stroke. In response to

statements on the TAQ, participants with stroke generally reported the device to be beneficial for their health, useful for monitoring their health, easy to use, and minimally intrusive. We observed generally agreeable responses to TAQ statements across all 7 dimensions of the instrument. Contrary to our hypothesis, more agreeable responses to the TAQ statements were not positively correlated with metrics of device wear, indicating that adherence to wearing the device should not be assumed even when participants report high device acceptance. In summary, this study provides new information about the acceptance of wearable activity monitors among persons with stroke and its association with real-world device wear.

## Funding

We acknowledge funding from the American Heart Association (grants 23IPA1054140 and 935556 to RTR and grant 24POST1187285 to GCB) and the Sheikh Khalifa Stroke Institute at Johns Hopkins Medicine.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Conflicts of Interest

None declared.

## References

1. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
2. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLOS Med 2016 Feb;13(2):e1001953. [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
3. Chang M, Büchel D, Reinecke K, Lehmann T, Baumeister J. Ecological validity in exercise neuroscience research: a systematic investigation. Eur J Neurosci 2022 Jan;55(2):487-509. [doi: 10.1111/ejn.15595] [Medline: 34997653]
4. Ingram JN, Wolpert DM. Naturalistic approaches to sensorimotor control. Prog Brain Res 2011;191:3-29. [doi: 10.1016/B978-0-444-53752-2.00016-3] [Medline: 21741541]
5. Peterson SM, Singh SH, Wang NXR, Rao RPN, Brunton BW. Behavioral and neural variability of naturalistic arm movements. eNeuro 2021;8(3):ENEURO.0007-21.2021. [doi: 10.1523/ENEURO.0007-21.2021] [Medline: 34031100]
6. International Classification of Functioning, Disability and Health (ICF). World Health Organization. 2001. URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2025-09-11]
7. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. The Lancet 2011 May;377(9778):1693-1702. [doi: 10.1016/S0140-6736(11)60325-5]
8. Lawrence ES, Coshall C, Dundas R, et al. Estimates of the prevalence of acute stroke impairments and disability in a multiethnic population. Stroke 2001 Jun;32(6):1279-1284. [doi: 10.1161/01.str.32.6.1279] [Medline: 11387487]
9. Li S. Stroke recovery is a journey: prediction and potentials of motor recovery after a stroke from a practical perspective. Life (Basel) 2023 Oct 15;13(10):2061. [doi: 10.3390/life13102061] [Medline: 37895442]
10. Fini NA, Holland AE, Keating J, Simek J, Bernhardt J. How is physical activity monitored in people following stroke? Disabil Rehabil 2015;37(19):1717-1731. [doi: 10.3109/09638288.2014.978508] [Medline: 25374044]
11. Ringeval M, Wagner G, Denford J, Paré G, Kitsiou S. Fitbit-based interventions for healthy lifestyle outcomes: systematic review and meta-analysis. J Med Internet Res 2020 Oct 12;22(10):e23954. [doi: 10.2196/23954] [Medline: 33044175]
12. McMahon SK, Lewis B, Oakes JM, Wyman JF, Guan W, Rothman AJ. Assessing the effects of interpersonal and intrapersonal behavior change strategies on physical activity in older adults: a factorial experiment. Ann Behav Med 2017 Jun;51(3):376-390. [doi: 10.1007/s12160-016-9863-z] [Medline: 28188585]
13. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173. [doi: 10.2196/mhealth.8211] [Medline: 29141837]
14. Walker K, Griffiths C, Yates J, Birkett-Swan L, Hina F, Ang CS. The experience and impact of a Fitbit-based physical activity and sleep hygiene intervention for older adults experiencing cognitive decline. AAR 2024;13(3):41-59. [doi: 10.4236/aar.2024.133004]
15. Wilson S, Beswick E, Morrell R, et al. Acceptance of wearable technology for the early detection of dementia‐causing diseases: perspectives from the CODEC II cohort. Alzheimer's & Dementia 2024 Dec;20(S2):e089597. [doi: 10.1002/alz.089597]
16. Hermsen S, Moons J, Kerkhof P, Wiekens C, De Groot M. Determinants for sustained use of an activity tracker: observational study. JMIR Mhealth Uhealth 2017 Oct 30;5(10):e164. [doi: 10.2196/mhealth.7311] [Medline: 29084709]
17. Lavelle G, Norris M, Flemming J, et al. Validity and acceptability of wearable devices for monitoring step-count and activity minutes among people with multiple sclerosis. Front Rehabil Sci 2021;2:737384. [doi: 10.3389/fresc.2021.737384] [Medline: 36188762]

18. Lee Y, Kozar KA, Larsen KRT. The technology acceptance model: past, present, and future. CAIS 2003;12. [doi: 10.17705/1CAIS.01250]
19. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
20. Chen K, Chan AHS. Predictors of gerontechnology acceptance by older Hong Kong Chinese. Technovation 2014 Feb;34(2):126-135. [doi: 10.1016/j.technovation.2013.09.010]
21. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 1;14(5):e133. [doi: 10.2196/jmir.2143] [Medline: 23026508]
22. Scherer R, Siddiq F, Tondeur J. The technology acceptance model (TAM): a meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. Comput Educ 2019 Jan;128:13-35. [doi: 10.1016/j.compedu.2018.09.009]
23. Demers M, Cain A, Bishop L, et al. Understanding stroke survivors' preferences regarding wearable sensor feedback on functional movement: a mixed-methods study. J NeuroEngineering Rehabil 2023 Nov 1;20(1). [doi: 10.1186/s12984-023-01271-z]
24. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 8;8(9):e18694. [doi: 10.2196/18694] [Medline: 32897239]
25. French MA, Keatley E, Li J, et al. The feasibility of remotely monitoring physical, cognitive, and psychosocial function in individuals with stroke or chronic obstructive pulmonary disease. Digit Health 2023;9:20552076231176160. [doi: 10.1177/20552076231176160] [Medline: 37214659]
26. Koffman LJ, Crainiceanu CM, Roemmich RT, French MA. Identifying unique subgroups of individuals with stroke using heart rate and steps to characterize physical activity. J Am Heart Assoc 2023 Sep 19;12(18):e030577. [doi: 10.1161/JAHA.123.030577] [Medline: 37681556]
27. Schriesheim CA, Hill KD. Controlling acquiescence response bias by item reversals: the effect on questionnaire validity. Educ Psychol Meas 1981 Dec;41(4):1101-1114. [doi: 10.1177/001316448104100420]
28. Van Domelen DR. Functions for processing accelerometer data. CRAN. 2018. URL: https://cran.r-project.org/web/packages/accelerometry/index.html [accessed 2025-12-11]
29. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. URL: https://www.r-project.org/ [accessed 2025-12-10]
30. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
31. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
32. de Looff PC, Noordzij ML, Nijman HLI, Goedhard L, Bogaerts S, Didden R. Putting the usability of wearable technology in forensic psychiatry to the test: a randomized crossover trial. Front Psychiatry 2024;15:1330993. [doi: 10.3389/fpsyt.2024.1330993] [Medline: 38947186]
33. Drehlich M, Naraine M, Rowe K, et al. Using the technology acceptance model to explore adolescents' perspectives on combining technologies for physical activity promotion within an intervention: usability study. J Med Internet Res 2020 Mar 6;22(3):e15552. [doi: 10.2196/15552] [Medline: 32141834]
34. Jiang A, Al-Dajani N, King C, Hong V, Koo HJ, Czyz E. Acceptability and feasibility of ecological momentary assessment with augmentation of passive sensor data in young adults at high risk for suicide. Psychiatry Res 2023 Aug;326:115347. [doi: 10.1016/j.psychres.2023.115347] [Medline: 37487460]
35. Michalsen H, Henriksen A, Pettersen G, et al. Using mobile health to encourage physical activity in individuals with intellectual disability: a pilot mixed methods feasibility study. Front Rehabil Sci 2023;4:1225641. [doi: 10.3389/fresc.2023.1225641] [Medline: 37691911]
36. Rosenberg D, Kadokura EA, Bouldin ED, Miyawaki CE, Higano CS, Hartzler AL. Acceptability of Fitbit for physical activity tracking within clinical care among men with prostate cancer. AMIA Annu Symp Proc 2016;2016:1050-1059. [Medline: 28269902]
37. Lang CE, Waddell KJ, Barth J, Holleran CL, Strube MJ, Bland MD. Upper limb performance in daily life approaches plateau around three to six weeks post-stroke. Neurorehabil Neural Repair 2021 Oct;35(10):903-914. [doi: 10.1177/15459683211041302] [Medline: 34510934]
38. Waddell KJ, Strube MJ, Tabak RG, Haire-Joshu D, Lang CE. Upper limb performance in daily life improves over the first 12 weeks poststroke. Neurorehabil Neural Repair 2019 Oct;33(10):836-847. [doi: 10.1177/1545968319868716] [Medline: 31431125]
39. Waddell KJ, Patel MS, Clark K, Harrington TO, Greysen SR. Effect of gamification with social incentives on daily steps after stroke: a randomized clinical trial. JAMA Neurol 2022 May 1;79(5):528-530. [doi: 10.1001/jamaneurol.2022.0231] [Medline: 35344027]

40. Waddell KJ, Patel MS, Wilkinson JR, et al. Deploying digital health technologies for remote physical activity monitoring of rural populations with chronic neurologic disease. Arch Rehabil Res Clin Transl 2023 Mar;5(1):100250. [doi: 10.1016/j.arrct.2022.100250] [Medline: 36968173]
41. Derungs A, Schuster-Amft C, Amft O. Physical activity comparison between body sides in hemiparetic patients using wearable motion sensors in free-living and therapy: a case series. Front Bioeng Biotechnol 2018;6:136. [doi: 10.3389/fbioe.2018.00136] [Medline: 30386777]
42. English C, Healy GN, Coates A, Lewis L, Olds T, Bernhardt J. Sitting and activity time in people with stroke. Phys Ther 2016 Feb;96(2):193-201. [doi: 10.2522/ptj.20140522] [Medline: 26112254]
43. Demers M, Bishop L, Cain A, et al. Wearable technology to capture arm use of stroke survivors in home and community settings: feasibility and early insights on motor performance. Phys Ther 2024 Feb 1;104(2):pzad172. [doi: 10.1093/ptj/pzad172] [Medline: 38166199]
44. Miller AE, Holleran CL, Bland MD, et al. Perspectives of key stakeholders on integrating wearable sensor technology into rehabilitation care: a mixed-methods analysis. Front Digit Health 2025;7:1534419. [doi: 10.3389/fdgth.2025.1534419] [Medline: 40357237]
45. Stock R, Gaarden AP, Langørgen E. The potential of wearable technology to support stroke survivors' motivation for home exercise - focus group discussions with stroke survivors and physiotherapists. Physiother Theory Pract 2024 Aug;40(8):1795-1806. [doi: 10.1080/09593985.2023.2217987] [Medline: 37246716]
46. French MA, Balasubramanian A, Hansel NN, et al. Impact of automated data flow and reminders on adherence and resource utilization for remotely monitoring physical activity in individuals with stroke or chronic obstructive pulmonary disease. medRxiv. 2024 Apr 18 p. 2024.04.15.24305852. [doi: 10.1101/2024.04.15.24305852] [Medline: 38699312]
47. Adams DA, Nelson RR, Todd PA. Perceived usefulness, ease of use, and usage of information technology: a replication. MIS Q 1992 Jun 1;16(2):227-247. [doi: 10.2307/249577]
48. Saini GS, Sharma S. Factors affecting consumers' perception towards e-payment systems in India. IJCS 2017 Jun 1;2(3):23. [doi: 10.17010/ijcs/2017/v2/i3/115039]
49. Bhattacherjee A. Understanding information systems continuance: an expectation-confirmation model. MIS Q 2001 Sep 1;25(3):351-370. [doi: 10.2307/3250921]
50. Liu Y, Park Y, Wang H. The mediating effect of user satisfaction and the moderated mediating effect of AI anxiety on the relationship between perceived usefulness and subscription payment intention. J Retail Consum Serv 2025 May;84:104176. [doi: 10.1016/j.jretconser.2024.104176]
51. Qi N, Koo JJ. A composite study of the factors influencing user subscription intention to artificial intelligence generated content - taking ChatGPT users as an example -. Korean Soc Sci Art 2024 Mar 31;42(2):457-476. [doi: 10.17548/ksaf.2024.03.30.457]
52. Hu PJ, Chau PYK, Sheng ORL, Tam KY. Examining the technology acceptance model using physician acceptance of telemedicine technology. J Manag Inf Syst 1999 Sep;16(2):91-112. [doi: 10.1080/07421222.1999.11518247]
53. Garavand A, Aslani N, Nadri H, Abedini S, Dehghan S. Acceptance of telemedicine technology among physicians: a systematic review. Inform Med Unlocked 2022;30:100943. [doi: 10.1016/j.imu.2022.100943]
54. Klingberg A, Sawe HR, Hammar U, Wallis LA, Hasselberg M. m-Health for burn injury consultations in a low-resource setting: an acceptability study among health care providers. Telemed J E Health 2020 Apr;26(4):395-405. [doi: 10.1089/tmj.2019.0048] [Medline: 31161967]
55. Almojaibel AA, Munk N, Goodfellow LT, et al. Health care practitioners' determinants of telerehabilitation acceptance. Int J Telerehab 2020 Jun 30;12(1):43-50. [doi: 10.5195/ijt.2020.6308]


## Abbreviations

**ICD-10:** International Classification of Diseases, Tenth Revision
**PEOU:** Perceived ease of use
**REDCap:** Research Electronic Data Capture
**TAQ:** Technology Acceptance Questionnaire


*Edited by A Coristine; submitted 02.05.25; peer-reviewed by A Hungbo, Y Shahsavar; revised version received 15.07.25; accepted 22.07.25; published 23.12.25.*

*Please cite as:*
*Nam J, Bellinger GC, Li J, French MA, T Roemmich R*
*Self-Reported Acceptance of a Wearable Activity Monitor in Persons With Stroke: Usability Study*
*JMIR Cardio 2025;9:e70007*
*URL: https://cardio.jmir.org/2025/1/e70007*
*doi:10.2196/70007*

# Toward Ambulatory Baroreflex Sensitivity: Comparison Between Indices of Arterial Line and Photoplethysmography in Male Volunteers

Jolanda Witteveen[1], MSc; Fabian Beutel[2], MSc, PhD; Evelien Hermeling[2], MSc, PhD

[1]One Planet Research Center, imec, Wageningen, The Netherlands

[2]imec The Netherlands/Holst Centre, High Tech Campus 31, Eindhoven, The Netherlands

**Corresponding Author:**
Evelien Hermeling, MSc, PhD
imec The Netherlands/Holst Centre, High Tech Campus 31, Eindhoven, The Netherlands

## Abstract

**Background:** Baroreflex sensitivity (BRS) is the body's ability to adjust heart rate (HR) to control blood pressure. Traditionally, BRS is quantified by measuring HR changes (obtained via an electrocardiogram [ECG]) following alterations in arterial pressure (conventionally measured through an arterial line). However, the invasiveness of arterial line necessitates alternatives, such as the volume clamp method and the less invasive pulse photoplethysmography (PPG). Notably, the PPG method is also suitable for continuous and free-living conditions.

**Objective:** This study aims to compare PPG-based features for BRS determination based on the volume clamp method and gold standard arterial line. Data from a previous study was used where the primary analysis focused on evaluating the accuracy of PPG-derived HR variability, while this analysis quantifies BRS by measuring HR changes following alterations in arterial line pressure or less invasive alternatives. In addition, we investigate the feasibility of assessing BRS patterns over 24 hours using data from a single volunteer.

**Methods:** A total of 28 male volunteers (age 52, SD 7 y; BMI 27, SD 4 kg/m$^2$) equipped with four sensing modalities: (1) arterial line [ABP], (2) infrared PPG, (3) volume clamp finger pressure (VCP), and (4) ECG, performed a protocol of 3 repetitive sessions in supine position. For the extended feasibility of continuous BRS measurement, ECG and PPG data were acquired for 24 hours in free-living conditions from a normotensive male volunteer (33 y). BRS index was calculated within the low-frequency window (0.04 - 0.15 Hz) averaging over all trials for each intervention and participant. A transfer function was estimated with systolic blood pressure (SBP) or its surrogate as input and HR (from the ECG) as output.

**Results:** PPG-based BRS features, specifically the rise-decay time ratio (RDRatio) and pulse arrival time (PAT), demonstrate intraparticipant precision of 44% and 23%, respectively, with interparticipant variation of 91% and 53%. The correlation of $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ with the gold standard $BRS_{SBP,ABP}$ (SBP) is 0.66 and 0.56, respectively. During intervention, the correlations remain high for $BRS_{RDRatio,PPG}$ (rest: 0.75, paced-breathing: 0.50, and handgrip: 0.46) and $BRS_{PAT,PPG}$ (rest: 0.69, paced-breathing: 0.52, and handgrip: 0.62). In the 24-hour data, the $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ exhibit changes during the day corresponding to the activity levels and SBP variations. Notably, during the night, a cyclic rhythm is observed for both $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$.

**Conclusions:** This study demonstrates that in male volunteers, PPG-based PAT and RDRatio BRS serve as suitable surrogates for gold-standard BRS derived from arterial line, showing the highest correlation and comparable intraparticipant coefficient variation. Furthermore, they show expected changes during controlled activities and a 24-hour feasibility test in free-living conditions.

***(JMIR Cardio 2025;9:e54771)*** doi:10.2196/54771

**KEYWORDS**

photoplethysmography; pulse arrival time; pulse transit time; blood pressure; pulse wave analysis; heart rate; heart rate variability; baroreflex; arterial line; circadian rhythm; heart; arterial; arterial line; feasibility; systolic blood pressure; cyclic rhythm; feasibility test; baroreceptor sensitivity

## Introduction

The baroreflex is a feedback system controlling arterial blood pressure (ABP). Stretch receptors in the aortic arch wall and carotid sinuses sense the changes in ABP. When arterial transmural pressure increases, the baroreflex responds by lowering heart rate (HR) and decreasing cardiac contractility and reducing peripheral vascular resistance, and vice versa [1].

Baroreflex sensitivity (BRS) refers to the body's ability to adjust HR in response to changes in blood pressure (BP). Maintaining this hemodynamic homeostasis is a continuous process and of vital importance to healthy organ perfusion. A decrease in baroreflex sensitivity is associated with (persistent) hypertension, heart failure, poor outcome after stroke and kidney failure [2-8].

BRS can be modeled in both the time and frequency domain. Traditionally, it is quantified by measuring the HR changes following variations in arterial pressure. The sequence method is a time-domain method in which 3 or more consecutive beats with progressively increasing or decreasing arterial pressure are followed by a progressive increase or decrease of HR [9]. The frequency domain or spectral analysis is applied on continuous electrocardiogram (ECG) and arterial pressure signals assuming changes in arterial pressure and HR, induced by the baroreflex, are oscillations in the same frequency band. In the frequency domain analysis, different strategies are used, including nonparametric transfer function, parametric transfer function, and the phase rectified signal averaging method [1,7]. The most used BRS method calculates the spectral relationship between the input signal, commonly the systolic blood pressure (SBP) obtained from the continuous arterial pressure signal, and the output signal, commonly the RR interval (interbeat time interval based on R-peak of ECG) obtained from the ECG [7,10]. The BRS is typically quantified as the average spectral gain within the low frequency (LF; 0.04 - 0.15 Hz) or high frequency (HF; 0.15 - 0.40 Hz) window [7,8,10,11].

The gold standard for measuring arterial pressure is directly through an arterial line, that is, a (fluid-filled) catheter inserted into an artery. However, due to the invasiveness of this technique, alternative methodologies are necessary. A commonly used alternative is the continuous finger BP measured using the volume clamp method. This method uses a 2-sensor system that combines photoplethysmography (PPG) and a pressure sensor [12,13]. Previous investigations have revealed that, depending on the device, the variability of the systolic pressure has been overestimated by 78% and 103% in the low-frequency band [13]. In addition, the same investigation demonstrated an underestimation of the baroreflex sensitivity by −24% or −31%. Another, more indirect method to estimate ABP is through PPG, currently predominantly used in research settings [14]. PPG measures the blood volume pulse through light transmission instead of a direct pressure pulse. The PPG signal is composed of a pulsatile ("AC" [alternating current]) and baseline ("DC" [direct current]) part. The AC part reflects the changes in blood volume and is divided into a systolic phase (from foot to primary peak) and a diastolic phase (from secondary peak to foot) [15,16]. The PPG volume pulse contour is related to a pressure pulse [17]. Primary peak amplitude, referred to as systolic amplitude, has been related to stroke volume and changes under the influence of vasomotor tone and blood volume [15,16]. Different features have been proposed relating PPG with BP, including pulse arrival time (PAT), pulse width, reflection index, and PPG variability [18]. PAT is the time between the electrical activation of the left ventricle, obtained with ECG, and the arrival of the wave in the periphery. PAT is known to be related to the BP or SBP [14]. As BP increases, the apparent arterial stiffness increases and PAT decreases. Besides this inverse relation, PAT is also determined by the pre-ejection period, that is, the time between electrical activation and opening of the valve of the left ventricle. PPG has also been related to systemic vascular resistance and vasomotor tone. The DC component of the PPG and pulse width are determined by, among other things, the vasomotor tone [16,19-21]. Pre-ejection period change, related to change in cardiac contractility, can also be derived from the PPG signal [22,23]. Hence, PPG contains more baroreflex-related information than just arterial pressure. It has been used to determine BRS and is most often compared to BRS, based on the volume clamp method [10,24-26]. To the best of our knowledge, BRS obtained from invasive arterial pressure, volume clamp finger pressure (VCP), and PPG have not been compared directly.

Continuous BRS measurements in a free-living condition could elucidate the variation of the BRS and its potential interaction with the circadian rhythm. BRS over a 24-hour period is useful to monitor autonomic nervous system (dys)function at night in the absence of other influences, to relate it to sleep stages, for example, in patients with prediabetes [27], older adults [28], or patients with hypertension [29,30]. Long-term free-living monitoring requires a minimally obtrusive wearable solution, which could be PPG, for instance. To reliably use PPG for BRS determination, it is important to understand the benefits and the disputes compared to a direct BP measurement from the arterial line.

This study aims to better understand which PPG-based features for BRS determination perform best in comparison to BRS based on the volume clamp method and gold standard arterial line. In addition, this study, in extension, also aims to assess the feasibility of assessing BRS patterns over 24 hours by means of a single volunteer.

## Methods

### Overview

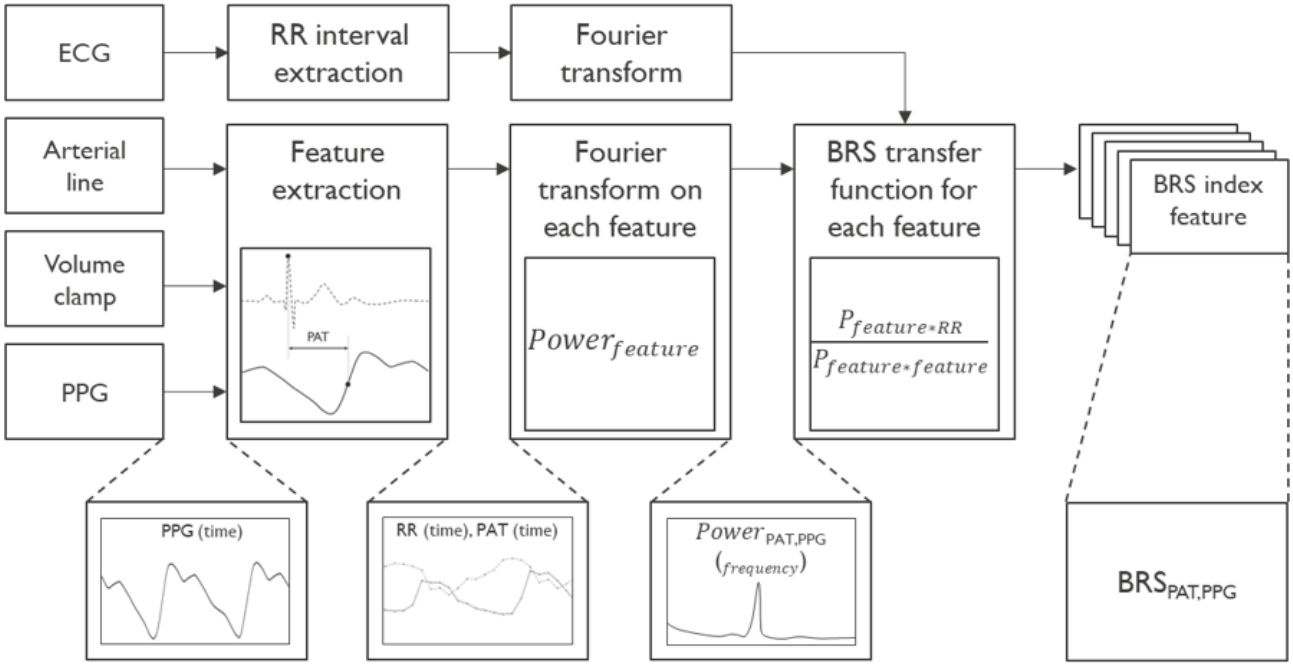
This study involves a secondary analysis of an existing dataset [31]. The primary analysis focused on evaluating the accuracy of PPG-derived HR variability (HRV) [31], while in this analysis, the baroreceptor sensitivity is quantified by measuring HR changes following alterations in arterial line pressure or less invasive alternatives.

### Datasets: Arterial Line Interventional Study

The interventional dataset was used to analyze the differences and similarities between BRS derived from invasive arterial line BP, VCP, and PPG. More details can be found in [31]. In summary, 28 male healthy volunteers (aged 52, SD 7 y; BMI

27, SD 4 kg/m$^2$, SBP 130, SD 12 mm Hg). Participants were equipped with four sensing modalities (see Figure 1): (1) arterial line inserted into radial artery on the nondominant arm, (2) infrared PPG at the index finger of the same arm (Biopac PPG100C, 240 Hz), (3) Finapress Nova at middle finger of the same arm, and (4) ECG in lead II configuration (Biopac ECG100C, 240 Hz). Participants performed a protocol of 3 repetitive sessions in supine position with the arm resting alongside the body. Each session included several interventions, namely two times paced breathing at 7 breaths for 3 minutes. A handgrip intervention during which the participant was asked to squeeze in a handgrip as much as possible for one and a half minutes using their dominant hand. In addition, a dedicated rest period where the participants were asked to close their eyes for 2 minutes. Extra (unlabeled) transition time was allocated in between activities.

**Figure 1.** Overview of the steps to compute baroreceptor sensitivity (BRS) indices with example for pulse arrival time (PAT) based on photoplethysmography (PPG). ECG: electrocardiogram; RR: interbeat time interval based on R-peak of ECG.




## Ambulatory 24-Hour Study

In addition, a single 24-hour recording on a healthy volunteer was used to demonstrate BRS trends under free living conditions. In a normotensive male (33 y), a wearable prototype developed by imec was placed that recorded ECG and PPG for 24 hours. The ECG was placed in lead II configuration, and the transmissive PPG sensor Nonin 8000J was placed on the left index finger. In addition, an ambulatory BP measurement device (Suntech Medical Oscar2) recorded cuff-based oscillometric BP from the left upper arm in intervals of 15 and 30 minutes during the day and night, respectively. This was a regular office day, including 8 hours of sleep, 2 walks, and office work behind a desk.

## Ethical Considerations

The interventional study dataset was collected at Ziekenhuis Oost-Limburg in Genk, Belgium, and has been approved by the institutional review board (ethical committee approval 16/039U). All enrolled participants were compensated with a US $135 voucher. The ambulatory 24-hour feasibility data have been approved by the institutional review board of imec The Netherlands in Eindhoven, the Netherlands.

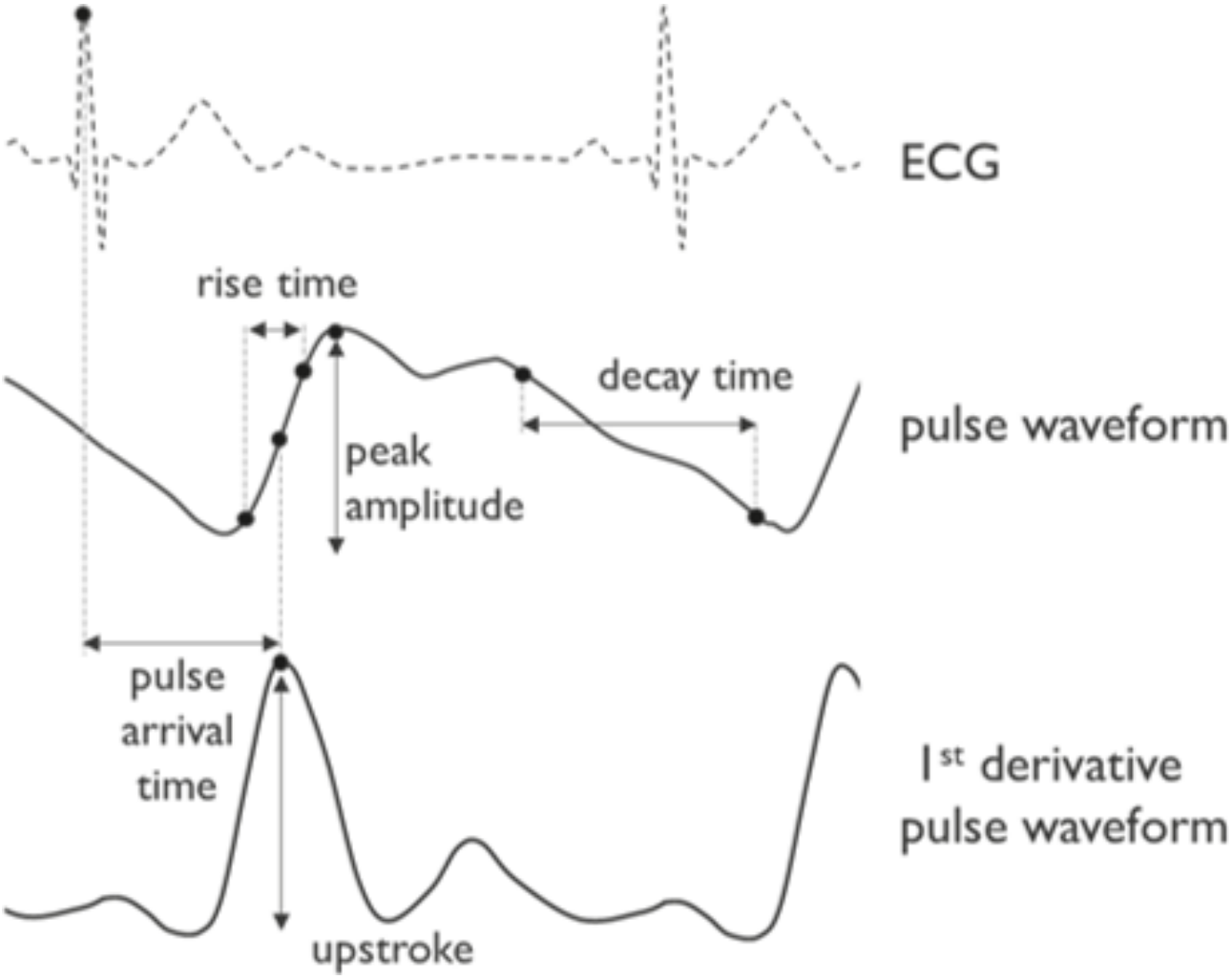
Both studies were conducted under the principles of the Declaration of Helsinki. All eligible participants were given the right to refuse participation and the right to withdraw from the study at any time. Written informed consent was collected from all participants before participation. To protect the participants' privacy, all data collected from this study were kept confidential and anonymized and were only accessible to the members of the research team.

## Data Analysis and Statistics

### *Data Preprocessing*

Data was processed and analyzed using Matlab R2022a (Mathworks). The relevant features in the pulse waveforms were computed from fiducial points detected in the first derivative signal, as described in detail by Fedjajevs et al [32] (see Figure 2). In brief, all data was band-pass filtered using a Butterworth low-pass filter with the cutoff at 10 Hz and high-pass filter acting as a differentiator. The differentiated signal is used to first find the upstroke, the local maximum of the differentiated signal. Next, the other fiducial points—peak, foot, shoulder, secondary peak, and dicrotic notch—were extracted using adaptive thresholding. Timestamps of the fiducial points are used to calculate the rise and decay times, amplitudes, and their ratios.

**Figure 2.** Visual representation of the features as described. Dashed line is the electrocardiogram (ECG) signal, the solid lines the pulse waveform from either photoplethysmography (PPG), volume clamp or arterial line, and first derivative of the PPG (PPG').




### *Features*

Features were computed per beat for all pulse waveforms, that is, PPG, volume clamp, and arterial line. Figure 2 shows a visual representation of the derived features. Peak amplitude (PA) is the amplitude of the first peak of the PPG. Pulse upstroke was defined as the amplitude of the peak in the first derivative of the PPG signal. Pulse arrival time (PAT) was defined as the delay between the R-peak from the ECG and the pulse upstroke in the PPG. Rise-Decay Ratio (RDRatio) was defined as the ratio between the rise time (10% to 70% of the peak amplitude in the systolic phase) and decay time (70% to 10% of the peak amplitude in the diastolic phase).

Volume clamp and arterial line features were derived in a similar manner as described for PPG and are indicated by the respective subscripts V (volume clamp) and A (ABP). SBP is the peak amplitude of either pressure pulse wave.

### *Transfer Function*

BRS indices of various modalities were calculated in the low frequency window (0.04 - 0.15 Hz) taking the average over all trials per intervention per participant. Data was resampled to 4 Hz, and a first-order trend was removed. A transfer function was estimated with feature (SBP or surrogate) as input and HR (RR interval derived from the ECG) as output. The transfer function was defined as the ratio of the cross power spectral density of the input (x, respectively BP) and output (y, respectively RR interval) and the spectral density of the input (x):

$$(1) H(f)=PxyPxx=Pfeature*RRPfeature*feature$$

Coherence levels were determined between SBP (or a surrogate feature) and RR interval for each BRS index. Considering the number of unique data points per segment (120 data points with 50% overlap), the threshold coherence at 95% CI is set at 0.14 [33,34].

### *Analysis of Arterial Line Data*

A structured analysis was done based on different BRS indices, calculated using the RR intervals from the ECG as input and as output the SBP from the arterial line ($BRS_{SBP,A}$ as gold reference) or other features as potential surrogates. The analysis included the following steps: Step 1, the direct comparison of BRS based on the SBP for the arterial line ($BRS_{SBP,ABP}$) and volume clamp ($BRS_{SBP,V}$) and the corresponding feature peak amplitude of the PPG ($BRS_{PeakAmp,PPG}$). Step 2, the comparison of various features (upstroke, PAT, and RDRatio) obtained from the arterial line sensor, as this is the signal that is the most direct BP recording with the highest signal to noise ratio. Followed by step 3, the comparison of BRS obtained from the same

features (SBP or PeakAmp, upstroke, PAT, and RDRatio) from arterial line and PPG. Finally, step 4 a comparison of BRS based on different features (SBP or PeakAmp, pulse upstroke, PAT, and RDRatio) obtained from volume clamp and PPG with the gold reference, namely BRS obtained from arterial line SBP.

### *Statistical Analysis*

For each participant and intervention, BRS measures were computed across all modalities and features. To mitigate the impact of outliers, we calculated the BRS index as the median value derived from three repeated measurements. The precision of each BRS index within each participant (intraparticipant precision) was determined by calculating the SD of the error. This error is the difference between the median of the participant and the individual values. The result was then expressed as a percentage of the mean BRS indices across all participants. Variation between participants (interparticipant variation) was calculated by determining the SD of the BRS indices. This was also expressed as a percentage of the mean BRS indices across all participants. In the final analysis, the correlation among the various BRS indices was determined using both Pearson and Spearman correlation methods, as not all BRS indices were normally distributed. It is important to note that the results from both the parametric (Pearson) and nonparametric (Spearman) methods were largely similar. Therefore, for simplicity, only the results derived from Pearson correlation analysis are reported. The level of statistical significance was set to .05.

### *Analysis 24-Hour Data*

The 24-hour dataset was processed in the same way as the interventional dataset with additional windowing. Based on results from arterial line data, PAT and RDRatio were selected as the best option for 24-hour BRS features (see Results and Discussion). From the filtered continuous ECG and PPG data acquired by the wearable prototype, beat-to-beat RR intervals, PAT, and RDRatio were computed. Both raw signals and extracted features were (dis)qualified based on an integration of 5 objective criteria:

First, any beat-to-beat samples 15 seconds before and 60 seconds post the cuff inflation due to occlusion.

Second, beat segments in the PPG signal were qualified using a proprietary signal-to-noise ratio (SNR) algorithm wherein a reference signal template is defined by 5 consecutive beats, and the noise impacting the signal morphology (eg, due to motion) is defined as the deviation of individual beats from this template and disqualified by empirical thresholds.

Third, absolute thresholding of beat-to-beat PAT samples deviating from a physiologically valid range, under the assumption of a fixed distance and pulse wave velocities from 2 to 10 m/seconds.

Fourth, variability thresholding of drastic beat-to-beat changes in HR and PAT deviating from physiologically valid ranges of HRV and PAT variability (respectively sympathetic changes in BP and arterial stiffness).

Fifth, disqualification of 1-minute epochs of persistent low quality, wherein the more robust ECG signal defines the expected number of cardiac cycles, and at least 50% of the PPG beats ought to be present and not undetected or disqualified by the previous criteria.

From here, $BRS_{PAT,P}$ and $BRS_{RDRatio,P}$ were calculated over a 4-minute sliding window (75% overlap), whereafter the median was taken from all BRS values exceeding the coherence threshold of 0.14 within a 1-hour sliding window (75% overlap). Cuff-based BP was measured every 15 minutes during the day and every 30 minutes during the night, providing at least two reference BP measurements included within the 1-hour sliding window. BP readings were qualified for validity by proprietary software of the ambulatory BP monitor (Suntech Medical Oscar2). The participant was also asked to stand still during the cuff inflation, which ensures stable ECG and PPG signals from the wearable system. For all clarity, no BRS was computed from the ambulatory cuff BP due to overlong sampling intervals. The BRS coherence was calculated over 4-minute windows per feature, and the median coherence for the subsequently computed 1-hour window held only the samples above the coherence threshold. Furthermore, given the longitudinal character of the 24-hour dataset, HRV as an indicator of autonomous nervous system activity was computed for relevant frequency bands: high frequency $HRV_{HF}$, reflecting the parasympathetic-driven respiratory band around 15 cycles per minute or 0.25 Hz on average, and the low frequency $HRV_{LF}$, reflecting baroreflex activity (balanced by sympathetic & parasympathetic activity) around 6 cycles per minute or 0.1 Hz on average [35]. Consistent with other features, the HRV indices were also averaged with a 1-hour sliding window. Ultimately, for visual inspection, all 1-hour averaged features are displayed on a time grid of 15-minute intervals.

## Results

### Arterial Line Intervention Study

The dataset consisted of 28 male volunteers. Per participant, 3 sessions were recorded of several interventions, including the interventions analyzed here: rest, paced breathing, and handgrip. A total of 420 segments were analyzed to extract BRS values of different (surrogate) features. By averaging over the repeated measures, we obtained 140 data points for each participant and each intervention.

For all participants and interventions, the mean of the $BRS_{SBP,ABP}$ and $BRS_{SBP,VCP}$ were 8.1 (SD 3.0) milliseconds/mm Hg and 6.6 (SD 3.0) milliseconds/mm Hg, respectively, with intraparticipant precision of 15% for $BRS_{SBP,ABP}$ and 21% for $BRS_{SBP,VCP}$, respectively. The PPG-based features show mean values of 3.2 (SD 1.9) au for $BRS_{PeakAmp,PPG}$, 93 (SD 60) au for $BRS_{Upstroke,PPG}$, 1.8 (SD 1.7) milliseconds for $BRS_{RDRatio,PPG}$, and 7.7 (SD 4.1) milliseconds/milliseconds for $BRS_{PAT,PPG}$.

The intraparticipant precision for the PPG-based features is on average higher: 54% for $BRS_{PeakAmp,PPG}$, 56% for $BRS_{Upstroke,PPG}$, and 44% for $BRS_{RDRatio,PPG}$. The intraparticipant precision of the $BRS_{PAT,PPG}$ approaches that of the traditional BRS index (23% for $BRS_{PAT,PPG}$, compared to 20% and 26% for $BRS_{SBP,ABP}$ and $BRS_{SBP,VCP}$, respectively). In all cases, the

interparticipant variation (37% for $BRS_{SBP,ABP,}$ 46% for $BRS_{SBP,VCP}$, 59% for $BRS_{PeakAmp,PPG}$, 64% for $BRS_{Upstroke,PPG}$, 91% for $BRS_{RDRatio,PPG}$, and 53% for $BRS_{PAT,PPG}$) exceeded the intraparticipant precision.

As described in the methods section, analysis was done in a step-by-step approach.

### *Step 1*

Direct comparison between the peak amplitude feature between sensor modalities showed that the correlation between $BRS_{SBP,ABP}$ and $BRS_{SBP,VCP}$ was 0.78. A lower correlation was found between the $BRS_{SBP,ABP}$ and $BRS_{PeakAmp,PPG}$ ($r$=0.59) and $BRS_{SBP,VCP}$ and $BRS_{PeakAmp,PPG}$ ($r$=0.56). These correlations were all highly significant. Note that the absolute values of the $BRS_{PeakAmp,PPG}$ cannot be compared to those of the SBP-based BRS values as the unit is different.

### *Step 2*

In Table 1, the correlation of BRS obtained from different features from the arterial line sensor showed that $BRS_{SBP,ABP}$ varied from strong to moderate for the different surrogate features $BRS_{RDRatio,ABP}$ ($r$=0.66), $BRS_{Upstroke,ABP}$ ($r$=0.54), and $BRS_{PAT,ABP}$ ($r$=0.46), all $P$<.05

**Table .** Correlation of baroreceptor sensitivity between alternative features derived from arterial line and systolic blood pressure from arterial line. Pearson's correlation between baroreceptor sensitivity index (BRS) of features extracted from arterial line with respect to the gold-standard reference feature systolic blood pressure from arterial line ($BRS_{SBP,ABP}$). For details about the features, see Figure 2 and the Methods section.

| Arterial line features | Correlation |
|---|---|
| $BRS_{PAT,ABP}$ | 0.46[a] |
| $BRS_{Upstroke,ABP}$ | 0.54[a] |
| $BRS_{RDRatio,ABP}$ | 0.66[a] |

[a]Indicates significant correlation ($P$<.05). $BRS_{PAT,ABP}$, $BRS_{Upstroke,ABP}$, and $BRS_{RDRatio,ABP}$ are BRS surrogate indices obtained from arterial derived features pulse arrival time, upstroke gradient, and rise time-decay time ratio, respectively.

### *Step 3*

Comparing the same feature between arterial line and PPG sensor modalities revealed that for peak-amplitude and upstroke derived BRS indices, the correlation was moderate and weak, respectively. In contrast, the correlation between PAT and RDRatio derived BRS indices across the sensor modalities was strong (see Table 2).

**Table .** Correlation between baroreflex sensitivity (BRS) extracted from arterial line (ABP) and PPG. BRS from arterial line (subscript $_{ABP}$) and photoplethysmography (PPG or subscript $_{PPG}$) using different features namely, systolic blood pressure (SBP), peak amplitude (PeakAmp), pulse arrival time (PAT), pulse upstroke (Upstroke) and rise time-decay time ratio (RDRatio). For details about the features, see Figure 2 and the Methods section. The primary peak (PeakAmp) of arterial line data is the systolic blood pressure (SBP).

| Arterial line | PPG[a] | Correlation (Arterial line vs PPG) |
|---|---|---|
| $BRS_{SBP,ABP}$ | $BRS_{PeakAmp,PPG}$ | 0.59[b] |
| $BRS_{PAT,ABP}$ | $BRS_{PAT,PPG}$ | 0.75[b] |
| $BRS_{Upstroke,ABP}$ | $BRS_{Upstroke,PPG}$ | 0.29[b] |
| $BRS_{RDRatio,ABP}$ | $BRS_{RDRatio,PPG}$ | 0.49[b] |

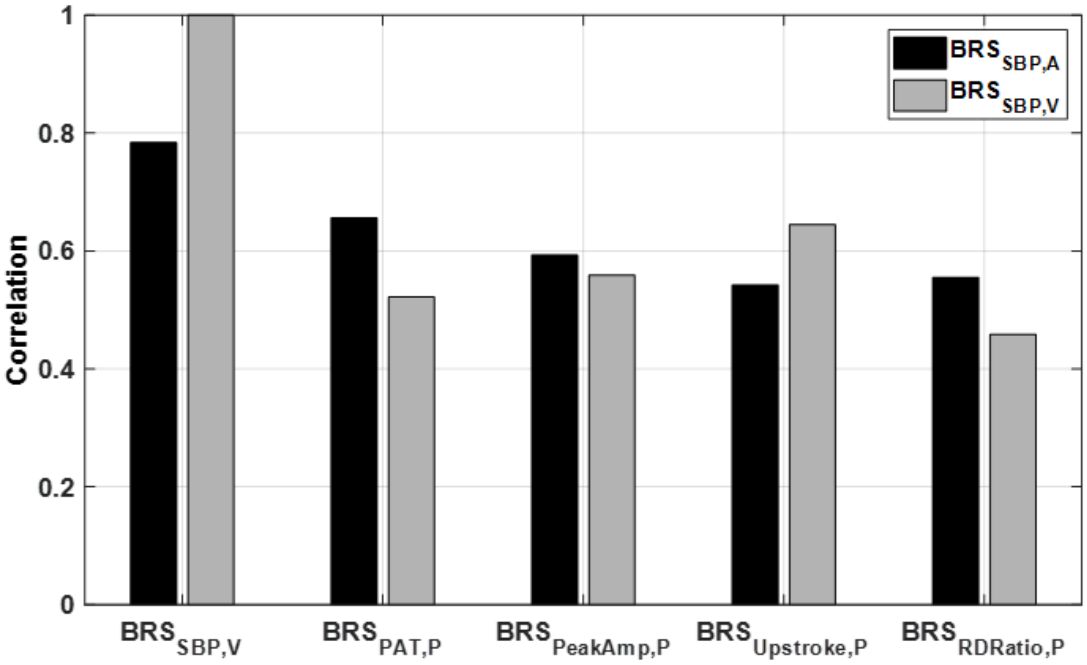
[a]PPG: photoplethysmography.

[b]Indicates significant correlation ($P$<.05).

### *Step 4*

Figure 3 shows the comparison of features derived from the PPG, as target sensor, and arterial line SBP, as gold standard or volume clamp. All correlations in step 4 were significant. As expected, the best correlation was observed between $BRS_{SBP,ABP}$ and $BRS_{SBP,VCP}$ ($r$=0.78). The PPG-based BRS surrogates had a strong to moderate correlation with $BRS_{SBP,ABP}$, where $BRS_{PAT,PPG}$ (0.66) had slightly higher values compared to $BRS_{PeakAmp,PPG}$, $BRS_{RDRatio,PPG}$, and $BRS_{Upstroke,PPG}$ (0.59, 0.56, and 0.54, respectively). However, when $BRS_{SBP,VCP}$ was used as an alternative reference to $BRS_{SBP,ABP}$ in the PPG-based surrogates that are based on PPG timing, a lower correlation was found with $BRS_{SBP,VCP}$ with respect to $BRS_{SBP,ABP}$ (0.52 vs 0.66 and 0.46 vs 0.56 for PAT and RDRatio, respectively). In contrast, the PPG amplitude derived parameters, that is, PeakAmp and upstroke had similar or even higher correlation with $BRS_{SBP,VCP}$ compared to $BRS_{SBP,ABP}$ (0.56 vs 0.59 and 0.64 vs 0.54, respectively; see Figure 3).

**Figure 3.** Correlation of baroreflex sensitivity (BRS) indices based on a selection of photoplethysmography (PPG) features with BRS from systolic blood pressure measured by arterial line ($BRS_{SBP,ABP}$, black) and the volume clamp method ($BRS_{SBP,VCP}$, gray). The BRS indices from PPG features $BRS_{PAT,PPG}$, $BRS_{PeakAmp,PPG}$, $BRS_{Upstroke,PPG}$, and $BRS_{RDRatio,PPG}$ are based on PPG and use pulse arrival time, peak amplitude, upstroke gradient, and rise time-decay time ratio, respectively. All correlations were significant ($P$<.05).




### *Interventions*

The gold-standard BRS, derived from arterial line SBP ($BRS_{SBP,ABP}$), exhibited an 11% increase during paced breathing and a 22% decrease during handgrip, compared to rest (see Figure 4). A similar pattern was observed for $BRS_{SBP,VCP}$ and $BRS_{PAT,PPG}$. However, the increase in BRS during paced breathing was more pronounced for these indices (58% and 42%, respectively). The $BRS_{RDRatio,PPG}$ showed a comparable increase (54%) during paced breathing as $BRS_{SBP,VCP}$ and $BRS_{PAT,PPG}$, but its reduction during handgrip was more significant (60%). The changes in BRS calculated from the other features $BRS_{PeakAmp,PPG}$ and $BRS_{Upstroke,PPG}$ were considerably larger during paced breathing with increases of 133% and 155% respectively. At rest, the correlation between $BRS_{SBP,ABP}$ and both $BRS_{SBP,VCP}$ and $BRS_{PAT,PPG}$ was similar. However, during paced breathing, the correlation with $BRS_{SBP,ABP}$ decreased for $BRS_{PAT,P}$ but increased for $BRS_{SBP,VCP}$ (see Table 3). The correlation between the other surrogate indices and the arterial line was generally lower across interventions, with the exception of $BRS_{RDRatio,PPG}$, which showed a high correlation with $BRS_{SBP,ABP}$ under rest conditions.

**Figure 4.** Baroreflex sensitivity (BRS) indices of different features for each intervention, namely: rest, paced breathing (pacedB), handgrip. $BRS_{SBP,ABP}$, $BRS_{SBP,VCP}$ are BRS indices based on systolic blood pressure measured by arterial line and volume clamp method, respectively. $BRS_{PAT,PPG}$, $BRS_{PeakAmp,PPG}$, $BRS_{Upstroke,PPG}$, $BRS_{RDRatio,PPG}$ are BRS indices measured using photoplethysmography using pulse arrival time, peak amplitude upstroke, and rise time-decay time ratio, respectively.

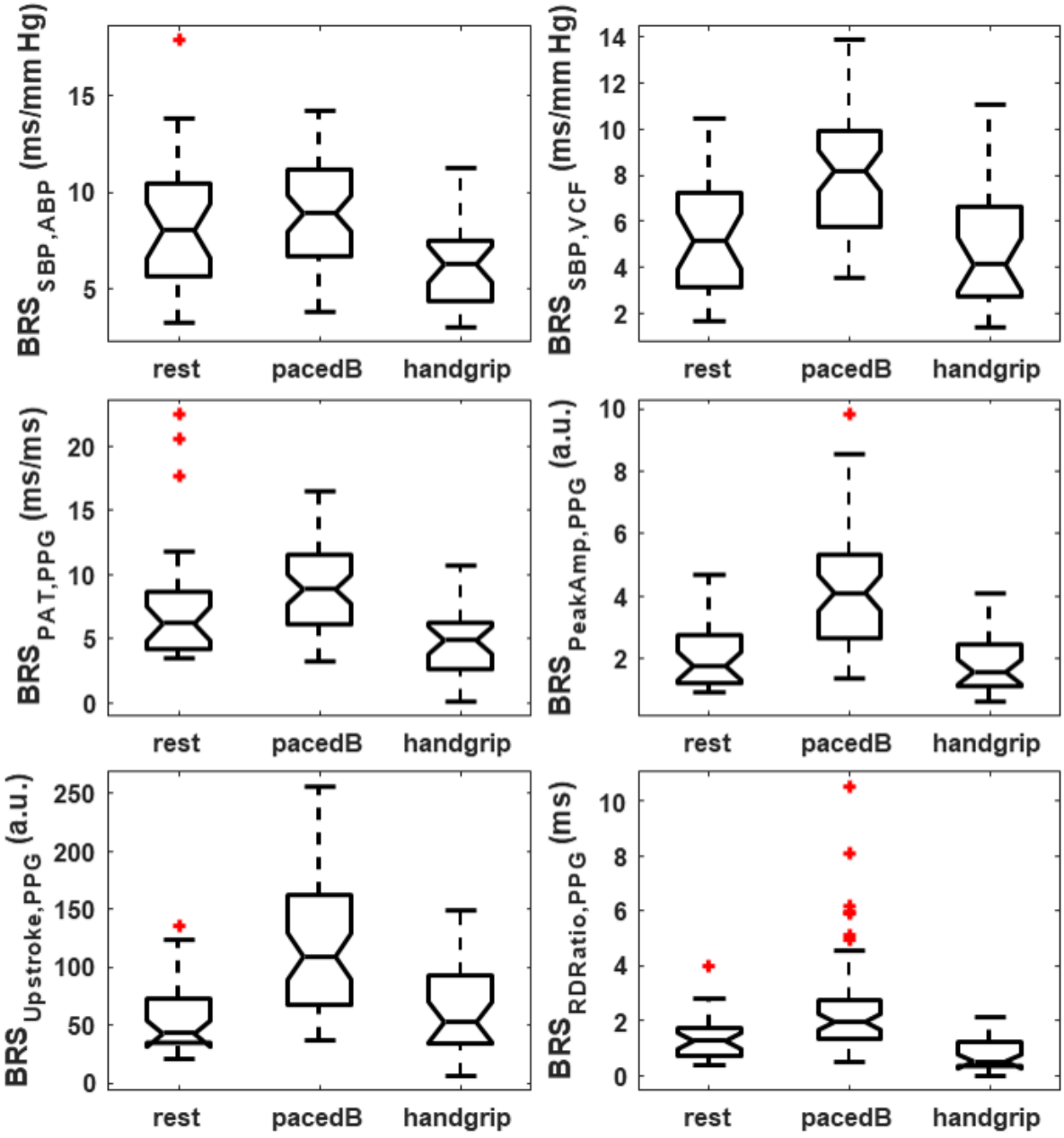


**Table .** Correlation between baroreflex sensitivity (BRS) extracted from volume clamp (VCP) and PPG and gold reference BRS based on arterial line derived systolic blood pressure ($BRS_{SBP,ABP}$). BRS obtained from systolic blood pressure measured by arterial line ($BRS_{SBP,ABP}$) is correlated to BRS obtained from SBP measured using volume clamp ($BRS_{SBP,VCP}$). In addition, BRS from photoplethysmography (PPG) is obtained with pulse arrival time (PAT), peak amplitude (PeakAmp), pulse upstroke (Upstroke), and rise time-decay time ratio (RDRatio). Note that the primary peak (PeakAmp) of arterial line and volume clamp data is the systolic blood pressure. Interventions are rest, paced breathing (pacedB), and handgrip. For further details, see the Methods section.

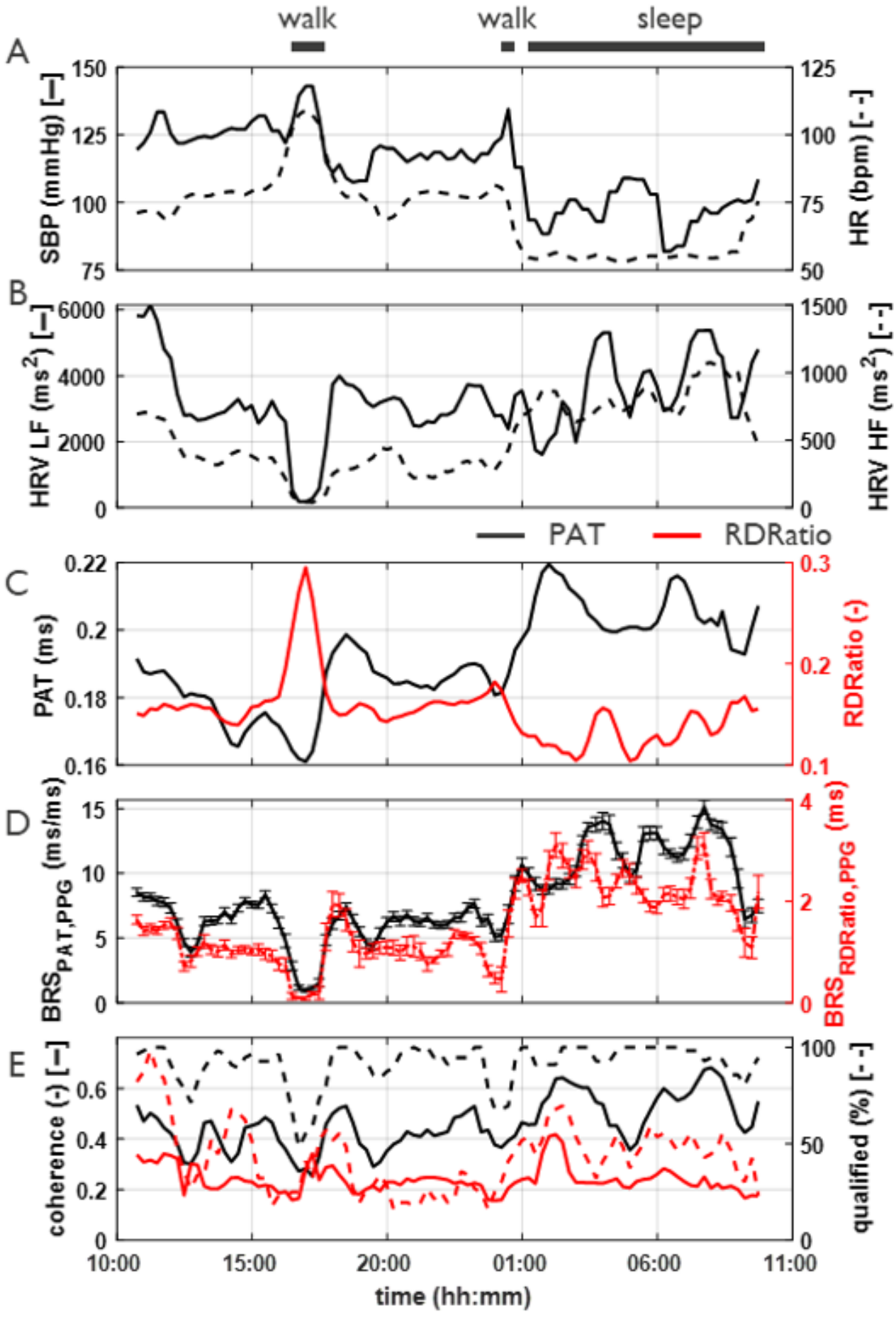
| Correlation to $BRS_{SBP,ABP}$ | Rest | PacedB | Handgrip |
|---|---|---|---|
| $BRS_{SBP,VCP}$ | 0.66[a] | 0.88[a] | 0.78[a] |
| $BRS_{PAT,PPG}$ | 0.69[a] | 0.52[a] | 0.62[a] |
| $BRS_{PeakAmp,PPG}$ | 0.56[a] | 0.51[a] | 0.53[a] |
| $BRS_{Upstroke,PPG}$ | 0.33 | 0.65[a] | 0.16 |
| $BRS_{RDRatio,PPG}$ | 0.75[a] | 0.50[a] | 0.46[a] |

[a]Indicates significant correlation ($P<.05$).

## Ambulatory 24-Hour Study

Figure 5 shows the exploratory 24-hour BRS recording from a healthy participant in free-living conditions. Except for a 1-hour walk around 5 PM and a short walk before 12 AM, the participant spent most of the day doing sedentary work. The participant was in bed between 12 and 8 AM. The lowest BRS was observed in the afternoon around 5 PM when the participant went for a walk, as also indicated by relatively high HR and low HRV. The activities were self-reported.

**Figure 5.** Analysis of the 24-hour trends for an individual participant. (A) trends of reference systolic blood pressure (based on upper arm cuff) and heart rate (HR) over time, (B) trends of low frequency (HRV LF) and high frequency heart rate variability (HRV HF). (C) Trends in the features pulse arrival time (PAT, in black) and rise time-decay time ratio (RDRatio, in red). (D) Trends in baroreflex sensitivity calculated from pulse arrival time ($BRS_{PAT,PPG}$) and rise time-decay time ratio ($BRS_{RDRatio,PPG}$). The error bars indicate the standard error of the BRS values within 15-minute segments. (E) Trends in coherence and qualified percentage of $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ values. The horizontal bars on top show the activities (walking and sleeping) of the individual.



The BRS indices show a clear pattern over the 24 hours, with both $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ being at a lower level during the day and increasing at night, effectively during sleep. Given the inverse relation between PAT and SBP, the mirrored pattern of increasing SBP and decreasing PAT (and vice versa) is clearly evident, also during walking activities with contributing HR. The correlation coefficients between PAT and RDRatio with SBP were high at −0.90 and −0.63 (both $P \leq .05$), respectively throughout the 24-hour trajectory. This confirms the essential validity of the observed trends based on the processed and qualified data. Interestingly, the correlation between SBP with the derived BRS indices ($BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$) was equally high and significant ($r=-0.74$ vs $r=-0.75$; both $P \leq .05$, respectively), while their mutual correlation was better ($r=0.84$; $P \leq .05$). At a close look, the rhythmic oscillations during the night cannot only be seen in $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ but also in $HRV_{LF}$. Unlike $HRV_{LF}$, $HRV_{HF}$ does not show any significant oscillations during the night, but instead displays a clear difference in absolute BRS level between day and night, which could not be observed in $HRV_{LF}$. Coherence of $BRS_{PAT,PPG}$ also shows a slight increase during the night. Except for a few datapoints during the walking activities, where the measurements were affected by motion artifacts, the percentage of $BRS_{PAT,PPG}$ above the coherence quality threshold of 0.14 never dropped below 80%, while the qualified coherence % of $BRS_{RDRatio,PPG}$ was significantly lower, as directly reflected in the coherence profile of $BRS_{RDRatio,PPG}$. At all times, even during the walking events where motion artifacts are to be expected, sufficient data is above coherence threshold ensuring valid median values throughout the day.

## Discussion

### Principal Findings

This study illustrates that both the BRS based on PAT and the RDRatio derived from PPG serve as appropriate substitutes for the gold-standard BRS obtained from arterial lines. This is substantiated by the highest correlation observed during rest, a comparable coefficient of variation within participants, as well as anticipated alterations during activities. Furthermore, the feasibility of these measures was successfully tested over a 24-hour period under free-living conditions. This underscores their potential applicability in real-world scenarios.

### Evaluation of BRS Index Surrogates

Baroreflex sensitivity based on arterial line SBP is best correlated with BRS determined from pulse arrival time derived from PPG (PAT,P) signal: 66% of the variation in $BRS_{SBP,ABP}$ is explained by $BRS_{PAT,PPG}$. The correlation of other surrogates to $BRS_{SBP,ABP}$ is slightly lower. In almost all cases, the interparticipant variation is higher than the intraparticipant precision, which suggests that these BRS features can be used to identify differences between individuals.

The baroreflex plays a crucial role in maintaining BP, acting through various pathways to modulate HR, vascular resistance, and cardiac contractility. The challenge in determining the BRS noninvasively is measuring (systolic) arterial pressure as the gold-standard method, arterial line, is invasive or obtrusive. The volume clamp method is an often-used noninvasive substitute to determine arterial pressure. It is known to overestimate the central ABP [36] and shown to overestimate BRS in the low frequencies [37]. Nevertheless, it has been used as a reference to validate other BRS methods, like those derived from PPG [10,24-26]. Comparing BRS using arterial pressure derived from the invasive arterial line with volume clamp method and PPG reveals the disputes and benefits of the methods. The BRS based on SBP derived from the volume clamp method shows the best correlation of 0.78 with BRS based on SBP from the invasive arterial line and a correlation of 0.59 with the BRS based on the systolic peak in the PPG signal. The latter is lower compared to results from others, reporting a correlation of 0.77 [10]. Nevertheless, our study population is considerably older (compared to 28.5 y) and has relatively high BP, which would lower the BRS and, in turn, increase the influence of noise, thereby reducing the correlation. An underestimation of the BRS based on SBP from the volume clamp method compared to the arterial line–based SBP of 24% was reported [13], similar to the 19% reported here.

The 3 modalities, arterial line, volume clamp, and PPG, have differences and similarities important to consider. The volume clamp method uses a PPG signal as well. Although the PPG signal is not used to measure the arterial pressure directly, it is used to maintain a constant volume by adjusting the cuff pressure, such that it equals the finger arterial pressure. Therefore, in contrast to the arterial line method, both PPG and volume clamp methods are influenced by peripheral perfusion and, hence, temperature, motion, and contact pressure. This could potentially cause errors when relating peripheral to systemic hemodynamic changes [15]. These errors would be visible between arterial line and volume clamp or PPG-derived features but be similarly present between PPG and volume clamp–derived features.

The range of BRS features derived from the PPG used here is also reported previously, like pulse upstroke, peak amplitude, pulse arrival time, and rise time [10,24]. In addition, models estimating SBP using PPG use PAT as the most important feature [14]. Pulse arrival time and rise-decay ratio from PPG correlate equally well with $BRS_{SBP,ABP}$ and $BRS_{SBP,VCP}$. Primary peak and pulse upstroke from PPG correlate well with $BRS_{SBP,VCP}$, but notably less with $BRS_{SBP,ABP}$, suggesting volume clamp underestimates the performance of $BRS_{PAT,P}$ and $BRS_{RdRatio,P}$ and overestimates the performance of others like $BRS_{Upstroke,P}$. These results suggest that $BRS_{SBP,V}$ can be used as a method to measure BRS in a noninvasive way; however, care should be taken to check the performance of other (PPG-derived) BRS indices using this method.

Coherence was used as a measure of BRS reliability. It assesses the similarity of two signals in the frequency domain; if the HR and arterial pressure have similar frequency components as a result of the baroreflex, the BRS becomes more reliable.

BRS changes from rest to controlled activities show how well a feature tracks the baroreflex effects of the interventions. The

correlation between SBP and HR increases for paced breathing intervention; therefore, an increase in BRS during paced breathing is expected. Especially at 6 breaths per minute, at which the HR and BP oscillation amplitude are increased [38]. Expectedly, during paced breathing, BRS increases for $BRS_{SBP,ABP}$ and all BRS surrogates compared to that at rest. However, the correlation between $BRS_{SBP,ABP}$ and $BRS_{PAT,PPG}$ decreases with paced breathing. It has been shown that during slow breathing of around 6 breaths/minute the BRS is overestimated, since, in this case, other mechanisms in which respiration influences HR also concentrate in the LF frequency band [39]. During the handgrip intervention, the coherence is increased, likely because of a thoracic pressure increase damping the oscillatory pressure effect of respiration (data not shown). An overestimation of the BRS changes based on volume clamp and PPG compared to $BRS_{SBP,ABP}$ is found, which is more prominent during paced breathing compared to handgrip or handgrip recovery.

### Ambulatory 24-Hour Study

The adequate robustness found for the BRS surrogates during controlled activities suggests a wider applicability for BRS monitoring, which was further assessed by means of a 24-hour recording under free-living conditions. Based on the structured analysis, the BRS based on PAT and RDRatio was considered the most promising to test for the 24-hour ambulatory; it had the lowest intraparticipant variation and highest correlation with $BRS_{SBP,ABP}$ during rest.

A wearable prototype for continuous ECG and PPG signal acquisition allowed for computation of beat-to-beat PAT, RDRatio, and RR intervals, and thereby enabled the observation of characteristic patterns in $BRS_{PAT,PPG}$, $BRS_{RDRatio,PPG}$, and HRV.

The BRS indices were found to be higher during the night as compared to daytime, which is in line with previous experiments where BRS was obtained from an arterial line [28,40,41]. This expected behavior of increasing nighttime BRS (and a coherence up to 0.6) could be explained by the absence of other inhibitory influences on the baroreflex like emotional behavior and somatic afferent influences stimulated by muscle contraction, as proposed by [40].

In preceding 24-hour BRS studies, oscillatory patterns of BRS during nighttime were less prominent [28,30,40], likely because these studies either average over participants or longer time windows and typically report one datapoint per hour (unlike the 15 min interval in this study). Meanwhile, studies focusing on sleep stages do show an increase in baseline BRS during the night and oscillations in BRS and $HRV_{LF}$ between rapid eye movement (REM) and non-REM sleep stages [42]. Supported by the findings from the interventional study and the coherence with $HRV_{LF}$, the observed patterns in the proposed BRS indices $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ undoubtedly reflect nocturnal BRS oscillations. Furthermore, the frequency of oscillations also appears to be in line with the typical duration and cycle times of adult sleep stages [43]. However, this remains to be further investigated with proper reference technology and in a larger population.

Furthermore, the proposed BRS indices may add value beyond existing HRV metrics. That is, both BRS indices show the baseline increase and nocturnal oscillations, whereas $HRV_{LF}$ does not show a clear baseline increase and $HRV_{HF}$ does not show oscillations, and its baseline increase may also be driven by different nocturnal respiration levels. Meanwhile, oscillations of $BRS_{PAT,PPG}$ tend to compare higher with $HRV_{LF}$; yet, $BRS_{RDRatio,PPG}$ seems to be more indicative of higher frequency contributions.

Regarding the reliability of the proposed indices, $BRS_{PAT,PPG}$ shows constantly high coherence throughout the 24 hours. Even during activity, a sufficient percentage of qualified samples is present, which may be further enhanced with basic signal and feature qualification strategies. Although the coherence of $BRS_{RDRatio,PPG}$ is substantially lower, it also remains above the threshold constantly with sufficient qualified samples. In terms of usability and technology integration, this may become a relevant compromise given that $BRS_{RDRatio,PPG}$ holds the theoretical advantage to be computed using a (peripheral) PPG, hence without an ECG.

### Clinical Implications

The nature of the PPG sensor also allows for free-living recordings, enabling the monitoring of the BRS of the patients on day-to-day activities. The obtained results from the 24-hour study are encouraging future research, considering the wide range of clinical applications for longitudinal BRS monitoring: as a prognostic tool for heart attacks and arrhythmias not only as single point measurement but also during sleep [11], and for cardiac mortality in patients with renal failure [44], or as a predictor of outcome after surgery [45]. Furthermore, present knowledge may be enhanced in day-to-day assessment of spontaneous BRS, which previously relied on 8-minute recordings on two consecutive days [46]. Variations in the 24-hour recordings between young and elderly people have also been reported [28]. Establishing a 24-hour recording could therefore show not only the BRS in short BP changes but also on circadian BP patterns. Ultimately, longitudinal BRS monitoring bears large potential for hypertension diagnostics, in particular for primary hypertension whose origin is widely unknown.

### Limitations

The study on arterial lines does present certain limitations, primarily due to the relatively limited sample size and the inclusion of only male participants. Further research is required to examine the impact of factors such as age and arterial stiffness on the BRS indices, as well as to explore their interrelationships. Such comprehensive analysis necessitates the involvement of larger and more diverse cohorts. Overall, these preliminary patterns of the $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ over a 24-hour period under free-living conditions support the findings from the controlled interventional study, demonstrating that it is possible to use PAT and RDRatio derived from the PPG signal to estimate the BRS. However, the key limitation of the 24-hour study is the confinement to a single participant, but the findings give rise to further expand this study to investigate circadian

and nocturnal BRS patterns in both healthy participants, and given the clinical relevance, also patient cohorts.

## Conclusions

BRS determined from pulse arrival time or RDRatio in a PPG signal is best correlated with BRS based on arterial line SBP. The $BRS_{PAT,PPG}$ and $BRS_{RDRatio,PPG}$ also follow the $BRS_{SBP,ABP}$ during different physical activities. Furthermore, it allows for 24-hour BRS recordings, in which the expected circadian rhythm patterns are present.

## Acknowledgments

This research is part of the Individualised Care from Early Risk of Cardiovascular Disease to Established Heart Failure (iCARE4CVD) project. iCARE4CVD has received funding from the Innovative Health Initiative Joint Undertaking (IHI JU) and BreakthroughT1D under grant agreement 101112022. The JU received support from the European Union's Horizon Europe research and innovation program and COCIR, EFPIA, Vaccines Europe, EuropaBio, and MedTech Europe. No generative AI was used in this manuscript.

## Data Availability

The data is owned by IMEC and the authors do not have permission to share the data publicly as we are bound to the European General Data Protection Regulation (GDPR) as well as the participants' consent, stating that data may only be used for specific purposes and not be shared with 3rd parties. This is because the dataset comprises personal identifiable data, which not only holds for demographics but also applies to electrocardiogram or arterial pressure waveforms. That being clarified, there may be ways to make anonymized or minimized data available on requests. However, this must be governed by a data sharing and/or processing agreement, which limits the use of the data (eg only to consented purpose, with no attempts to re-identify participants, etc.). Please contact privacy@imec.nl.

## Authors' Contributions

EH contributed to conceptualization. JW and FB performed formal analysis. JW, FB, and EH contributed to methodology. JW and FB handled the software. EH performed supervision. JW and FB contributed to writing the original draft. EH contributed to writing, review, and editing.

## Conflicts of Interest

None declared.

## References

1. La Rovere MT, Pinna GD, Raczak G. Baroreflex sensitivity: measurement and clinical implications. Ann Noninvasive Electrocardiol 2008 Apr;13(2):191-207. [doi: 10.1111/j.1542-474X.2008.00219.x] [Medline: 18426445]
2. Larson M, Chantigian DP, Asirvatham-Jeyaraj N, Van de Winckel A, Keller-Ross ML. Slow-paced breathing and autonomic function in people post-stroke. Front Physiol 2020;11:573325. [doi: 10.3389/fphys.2020.573325] [Medline: 33192570]
3. Lin CH, Yen CC, Hsu YT, et al. Baroreceptor sensitivity predicts functional outcome and complications after acute ischemic stroke. J Clin Med 2019 Mar 3;8(3):300. [doi: 10.3390/jcm8030300] [Medline: 30832391]
4. Robinson TG, Dawson SL, Eames PJ, Panerai RB, Potter JF. Cardiac baroreceptor sensitivity predicts long-term outcome after acute ischemic stroke. Stroke 2003 Mar;34(3):705-712. [doi: 10.1161/01.STR.0000058493.94875.9F] [Medline: 12624295]
5. Tang S, Xiong L, Fan Y, Mok VCT, Wong KS, Leung TW. Stroke outcome prediction by blood pressure variability, heart rate variability, and baroreflex sensitivity. Stroke 2020 Apr;51(4):1317-1320. [doi: 10.1161/STROKEAHA.119.027981] [Medline: 31964286]
6. Ormezzano O, Cracowski JL, Quesada JL, Pierre H, Mallion JM, Baguet JP. EVAluation of the prognostic value of BARoreflex sensitivity in hypertensive patients: the EVABAR study. J Hypertens (Los Angel) 2008;26(7):1373-1378. [doi: 10.1097/HJH.0b013e3283015e5a]
7. Pinna GD, Porta A, Maestri R, De Maria B, Dalla Vecchia LA, La Rovere MT. Different estimation methods of spontaneous baroreflex sensitivity have different predictive value in heart failure patients. J Hypertens (Los Angel) 2017;35(8):1666-1675. [doi: 10.1097/HJH.0000000000001377]
8. Bari V, Vaini E, Pistuddi V, et al. Comparison of causal and non-causal strategies for the assessment of baroreflex sensitivity in predicting acute kidney dysfunction after coronary artery bypass grafting. Front Physiol 2019;10(October):1319. [doi: 10.3389/fphys.2019.01319] [Medline: 31681021]
9. Silva LEV, Dias DPM, da Silva CAA, Salgado HC, Fazan R Jr. Revisiting the sequence method for baroreflex analysis. Front Neurosci 2019;13(JAN):17. [doi: 10.3389/fnins.2019.00017] [Medline: 30728765]

10. Lázaro J, Gil E, Orini M, Laguna P, Bailón R. Baroreflex sensitivity measured by pulse photoplethysmography. Front Neurosci 2019;13(April):339. [doi: 10.3389/fnins.2019.00339] [Medline: 31057351]
11. Swenne CA. Baroreflex sensitivity: mechanisms and measurement. Neth Heart J 2013 Feb;21(2):58-60. [doi: 10.1007/s12471-012-0346-y] [Medline: 23179611]
12. Langewouters GJ, Settels JJ, Roelandt R, Wesseling KH. Why use Finapres or Portapres rather than intra-arterial or intermittent non-invasive techniques of blood pressure measurement? J Med Eng Technol 1998;22(1):37-43. [doi: 10.3109/03091909809009997] [Medline: 9491357]
13. Maestri R, Pinna GD, Robbi E, Capomolla S, La Rovere MT. Noninvasive measurement of blood pressure variability: accuracy of the Finometer monitor and comparison with the Finapres device. Physiol Meas 2005 Dec;26(6):1125-1136. [doi: 10.1088/0967-3334/26/6/021] [Medline: 16311459]
14. Elgendi M, Fletcher R, Liang Y, et al. The use of photoplethysmography for assessing hypertension. npj Digit Med 2019 Dec 1;2(1). [doi: 10.1038/s41746-019-0136-7]
15. Elgendi M. On the analysis of fingertip photoplethysmogram signals. Curr Cardiol Rev 2012 Feb;8(1):14-25. [doi: 10.2174/157340312801215782] [Medline: 22845812]
16. Castaneda D, Esparza A, Ghamari M, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. Int J Biosens Bioelectron 2018;4(4):195-202. [doi: 10.15406/ijbsbe.2018.04.00125] [Medline: 30906922]
17. Millasseau SC, Guigui FG, Kelly RP, et al. Noninvasive assessment of the digital volume pulse. Hypertension 2000 Dec;36(6):952-956. [doi: 10.1161/01.HYP.36.6.952]
18. Sun S, Bezemer R, Long X, Muehlsteff J, Aarts RM. Systolic blood pressure estimation using PPG and ECG during physical exercise. Physiol Meas 2016 Dec 1;37(12):2154-2169. [doi: 10.1088/0967-3334/37/12/2154]
19. Awad AA, Haddadin AS, Tantawy H, et al. The relationship between the photoplethysmographic waveform and systemic vascular resistance. J Clin Monit Comput 2007 Dec;21(6):365-372. [doi: 10.1007/s10877-007-9097-5] [Medline: 17940842]
20. Gurel NZ, Huang M, Wittbrodt MT, et al. Quantifying acute physiological biomarkers of transcutaneous cervical vagal nerve stimulation in the context of psychological stress. Brain Stimul 2020;13(1):47-59. [doi: 10.1016/j.brs.2019.08.002] [Medline: 31439323]
21. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 2007 Mar;28(3):R1-39. [doi: 10.1088/0967-3334/28/3/R01] [Medline: 17322588]
22. Beutel F, Van Hoof C, Rottenberg X, Reesink K, Hermeling E. Pulse arrival time segmentation into cardiac and vascular intervals - implications for pulse wave velocity and blood pressure estimation. IEEE Trans Biomed Eng 2021 Sep;68(9):2810-2820. [doi: 10.1109/TBME.2021.3055154] [Medline: 33513094]
23. Reesink KD, Hermeling E, Hoeberigs MC, Reneman RS, Hoeks APG. Carotid artery pulse wave time characteristics to quantify ventriculoarterial responses to orthostatic challenge. J Appl Physiol (1985) 2007 Jun;102(6):2128-2134. [doi: 10.1152/japplphysiol.01206.2006] [Medline: 17317873]
24. Abe M, Yoshizawa M, Obara K, Sugita N, Homma N, Yambe T. Evaluation of baroreflex function using green light photoplethysmogram in consideration of resistance to artifacts. ABE 2015;4:1-6. [doi: 10.14326/abe.4.1]
25. Martínez-García P, Lerma C, Infante Ó. Relation of the baroreflex mechanism with the photoplethysmographic volume in healthy humans during orthostatism. Arch Cardiol Mex 2012;82(2):82-90. [Medline: 22735647]
26. Mol A, Meskers CGM, Niehof SP, Maier AB, van Wezel RJA. Pulse transit time as a proxy for vasoconstriction in younger and older adults. Exp Gerontol 2020 Jul 1;135(April):110938. [doi: 10.1016/j.exger.2020.110938] [Medline: 32247853]
27. Wang W, Redline S, Khoo MCK. Autonomic markers of impaired glucose metabolism: effects of sleep-disordered breathing. J Diabetes Sci Technol 2012 Sep 1;6(5):1159-1171. [doi: 10.1177/193229681200600521] [Medline: 23063043]
28. Parati G, Frattola A, Di Rienzo M, Castiglioni P, Pedotti A, Mancia G. Effects of aging on 24-h dynamic baroreceptor control of heart rate in ambulant subjects. Am J Physiol 1995 Apr;268(4 Pt 2):H1606-H1612. [doi: 10.1152/ajpheart.1995.268.4.H1606] [Medline: 7733361]
29. Khoo MCK, Wang W, Chalacheva P. Monitoring ultradian changes in cardiorespiratory control in obstructive sleep apnea syndrome. Annu Int Conf IEEE Eng Med Biol Soc 2011;2011:1487-1490. [doi: 10.1109/IEMBS.2011.6090363] [Medline: 22254601]
30. Vaile JC, Stallard TJ, Al-Ani M, Jordan PJ, Townend JN, Littler WA. Sleep and blood pressure: spontaneous baroreflex sensitivity in dippers and non-dippers. J Hypertens (Los Angel) 1996 Dec;14(12):1427-1432. [doi: 10.1097/00004872-199612000-00007]
31. Fedjajevs A, Groenendaal W, Grieten L, Agell C, Vandervoort PM, Hermeling E. Evaluation of HRV from repeated measurements of PPG and arterial blood pressure signals. Presented at: 2021 Computing in Cardiology (CinC); Sep 13-15, 2021; Brno, Czech Republic p. 48-51. [doi: 10.23919/CinC53138.2021.9662673]
32. Fedjajevs A, Groenendaal W, Agell C, Hermeling E. Platform for analysis and labeling of medical time series. Sensors (Basel) 2020 Dec 19;20(24):7302. [doi: 10.3390/s20247302] [Medline: 33352643]
33. Saeed NP, Reneman RS, Hoeks APG. Contribution of vascular and neural segments to baroreflex sensitivity in response to postural stress. J Vasc Res 2009;46(5):469-477. [doi: 10.1159/000200962] [Medline: 19204404]

34. Wang SY, Liu X, Yianni J, Christopher Miall R, Aziz TZ, Stein JF. Optimising coherence estimation to assess the functional correlation of tremor-related activity between the subthalamic nucleus and the forearm muscles. J Neurosci Methods 2004 Jul 30;136(2):197-205. [doi: 10.1016/j.jneumeth.2004.01.008] [Medline: 15183272]
35. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health 2017;5:258. [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]
36. Waddell TK, Dart AM, Medley TL, Cameron JD, Kingwell BA. Carotid pressure is a better predictor of coronary artery disease severity than brachial pressure. Hypertension 2001 Oct;38(4):927-931. [doi: 10.1161/hy1001.096107] [Medline: 11641311]
37. Lombardi F, Parati G. An update on: cardiovascular and respiratory changes during sleep in normal and hypertensive subjects. Cardiovasc Res 2000 Jan 1;45(1):200-211. [doi: 10.1016/s0008-6363(99)00329-6] [Medline: 10728336]
38. Russo MA, Santarelli DM, O'Rourke D. The physiological effects of slow breathing in the healthy human. Breathe (Sheff) 2017 Dec;13(4):298-309. [doi: 10.1183/20734735.009817] [Medline: 29209423]
39. Frederiks J, Swenne CA, TenVoorde BJ, et al. The importance of high-frequency paced breathing in spectral baroreflex sensitivity assessment. J Hypertens (Los Angel) 2000 Nov;18(11):1635-1644 [FREE Full text] [doi: 10.1097/00004872-200018110-00015]
40. Di Rienzo M, Parati G, Castiglioni P, Tordi R, Mancia G, Pedotti A. Baroreflex effectiveness index: an additional measure of baroreflex control of heart rate in daily life. Am J Physiol Regul Integr Comp Physiol 2001 Mar;280(3):R744-R751. [doi: 10.1152/ajpregu.2001.280.3.R744] [Medline: 11171653]
41. Persson PB, DiRienzo M, Castiglioni P, et al. Time versus frequency domain techniques for assessing baroreflex sensitivity. J Hypertens (Los Angel) 2001 Oct;19(10):1699-1705. [doi: 10.1097/00004872-200110000-00001]
42. Legramante JM, Marciani MG, Placidi F, et al. Sleep-related changes in baroreflex sensitivity and cardiovascular autonomic modulation. J Hypertens (Los Angel) 2003 Aug;21(8):1555-1561. [doi: 10.1097/00004872-200308000-00021]
43. Patel A, Reddy V, Shumway K. Physiology, sleep stages. In: StatPearls 2022.
44. Johansson M, Gao SA, Friberg P, et al. Baroreflex effectiveness index and baroreflex sensitivity predict all-cause mortality and sudden death in hypertensive patients with chronic renal failure. J Hypertens (Los Angel) 2007;25(1):163-168 [FREE Full text] [doi: 10.1097/01.hjh.0000254377.18983.eb]
45. Ranucci M, Porta A, Bari V, Pistuddi V, La Rovere MT. Baroreflex sensitivity and outcomes following coronary surgery. In: Bianchi C, editor. PLoS ONE 2017;12(4):e0175008. [doi: 10.1371/journal.pone.0175008] [Medline: 28384188]
46. Maestri R, Raczak G, Torunski A, et al. Day-by-day variability of spontaneous baroreflex sensitivity measurements: implications for their reliability in clinical and research applications. J Hypertens (Los Angel) 2009;27(4):806-812. [doi: 10.1097/HJH.0b013e328322fe4b]


## Abbreviations

**ABP:** arterial line blood pressure
**AC:** alternating current
**BP:** blood pressure
**BRS:** baroreceptor sensitivity
**DC:** direct current
**ECG:** electrocardiogram
**HF:** high frequency
**HR:** heart rate
**HRV:** heart rate variability
**LF:** low frequency
**PA:** peak amplitude
**PAT:** pulse arrival time
**PPG:** photoplethysmography
**RDRatio:** rise-time decay-time ratio
**REM:** rapid eye movement
**RR interval:** interbeat time interval based on R-peak of ECG
**SBP:** systolic blood pressure
**SNR:** signal to noise ratio
**VCP:** volume clamp finger pressure

*Edited by A Coristine; submitted 21.11.23; peer-reviewed by H Liu, MK Skoric, S Thirunavukkarasu, X Xing; revised version received 18.04.25; accepted 18.04.25; published 17.07.25.*

*Please cite as:*
*Witteveen J, Beutel F, Hermeling E*
*Toward Ambulatory Baroreflex Sensitivity: Comparison Between Indices of Arterial Line and Photoplethysmography in Male Volunteers*
*JMIR Cardio 2025;9:e54771*
*URL: https://cardio.jmir.org/2025/1/e54771*
*doi:10.2196/54771*

# A Medication Management App (Smart-Meds) for Patients After an Acute Coronary Syndrome: Pilot Pre-Post Mixed Methods Study

Frederic Ehrler[1], PhD; Liliane Gschwind[2], PhD; Hamdi Hagberg[1], MS; Philippe Meyer[3,4], MD, PhD; Katherine Blondon[4,5], MD, PhD

[1]Information Systems Directorate, University Hospital of Geneva, Geneva, Switzerland

[2]Department of Pharmacy, University Hospital of Geneva, Geneva, Switzerland

[3]Service of Cardiology, University Hospital of Geneva, Geneva, Switzerland

[4]Medicine Faculty, University of Geneva, Geneva, Switzerland

[5]Medical Directorate, University Hospital of Geneva, Geneva, Switzerland

**Corresponding Author:**
Frederic Ehrler, PhD
Information Systems Directorate, University Hospital of Geneva, Geneva, Switzerland

## Abstract

**Background:** Medication nonadherence remains a significant challenge in the management of chronic conditions, often leading to suboptimal treatment outcomes and increased health care costs. Innovative interventions that address the underlying factors contributing to nonadherence are needed. Gamified mobile apps have shown promise in promoting behavior change and engagement.

**Objective:** This pilot study aimed to evaluate the efficacy and usability of a gamified mobile app that used a narrative storytelling approach to enhance medication adherence among patients following acute coronary syndrome (ACS). The study aimed to assess changes in participants' beliefs about medication and self-reported adherence before and after the intervention. Additionally, user feedback regarding the narrative component of the app was gathered.

**Methods:** Overall, 18 patients who recently experienced ACS were recruited for a 1-month intervention using the gamified app. Participants' beliefs about medication and self-reported adherence were assessed using standardized scales pre- and postintervention. The app's usability was also evaluated through a postintervention questionnaire. Statistical analyses were performed to determine the significance of changes in belief and adherence scores.

**Results:** Although 33% (6/18) of the participants did not use the intervention more than once, the remaining 12 remained engaged during the 30 days of the study. The results did not indicate a significant improvement in participants' beliefs about medication following the intervention. However, self-reported adherence significantly improved ($P$<.05) after the intervention with a mean score going from 29.1 (SD 6.9) to 32.4 (SD 5.6), with participants demonstrating a greater self-efficacy to their prescribed medication regimen. However, the results did not indicate a significant improvement in participants' beliefs about medication. With a mean average score of 80.6, the usability evaluation indicates a good usability rating for the gamified app. However, the narrative storytelling component of the app was not favored by the participants, as indicated by their feedback.

**Conclusions:** This pilot study suggests that a gamified mobile app using narration may effectively enhance medication self-efficacy and positively influence patients' beliefs about medication following ACS. However, the narrative component of the app did not receive favorable feedback from participants. Future research should focus on exploring alternative methods to engage participants in the app's narrative elements while maintaining the positive impact on adherence and beliefs about medication observed in this study.

***(JMIR Cardio 2025;9:e50693)*** doi:10.2196/50693

**KEYWORDS**

medication adherence; gamified app; narration; acute coronary syndrome; beliefs about medication; self-reported adherence; pilot study; usability evaluation; storytelling component

## Introduction

Medication nonadherence is a well-identified health care issue, particularly for chronic diseases. Poor adherence worsens clinical outcomes and induces higher downstream rehospitalization rates as well as a higher use of resources [1]. Despite the physicians' efforts to convey the importance of the medications they prescribe, patients still find several intentional or unintentional reasons for deviating from their treatment plan [2]. Prior research reports that the most common factors associated with nonadherence are forgetfulness (50%), having other medications to take (20%), and being symptom-free (20%) [3]. The risk of poor adhesion is further increased with the medication regimen complexity, which increases with each decision about taking medication that a patient needs to make [4].

After an acute coronary syndrome (ACS), secondary cardiovascular prevention recommendations mainly involve lifestyle changes (eg, physical activity, smoking, or diet) and adherence to the prescribed drug regimen [5]. Patients with ACS are at particular risk of failing to adhere to their medication regimen since they may lack comprehension of medication importance, and have difficulty accessing medication, or affording the medication [6]. Additionally, medications used to treat ACS can have significant side effects that can make it difficult to take them regularly [7]. Patients with ACS may also need to take multiple medications, and there is a risk of drug interactions between them [8]. Moreover, the various medications used to treat ACS require regular monitoring to ensure they are working properly and to monitor the side effects. Finally, the medications used to treat ACS often require a longer time, which can be difficult for some patients to adhere to [9].

Mobile health apps provide new opportunities to support medication adherence [10]. First, they can remind users to take their medication on time. This can help ensure that users do not forget to take their medication or take incorrect doses. For instance, a meta-analysis of SMS text messaging interventions to improve adherence to medication in chronic diseases showed that SMS text message reminders were associated with increased odds of being adherent [11]. Second, mobile apps can track patients' medication use and provide feedback on their progress. They can offer personalized advice for treatment and behavioral change support, as well as facilitate communication between patients and their health care professionals [12]. This can help patients keep track of their medication use and identify any issues that may be preventing them from taking their medication as prescribed. Finally, mobile apps can connect users with health care professionals and support groups to provide additional motivation and help. This can help patients stay on track with their medication use and provide emotional support when needed.

Gamification for health behavior change involves applying game design elements and principles to encourage and motivate individuals to adopt healthier behaviors. It leverages techniques such as rewards, challenges, competition, and progress tracking to engage users in activities that promote better health outcomes. Examples include fitness apps that award points for completing workouts, digital platforms that encourage healthy eating through virtual rewards, and wearable devices that gamify physical activity by setting goals and providing feedback. By making health-related tasks more enjoyable and interactive, gamification aims to increase user motivation, adherence to health goals, and overall well-being [13]. Gamification is a mechanism that has proven to be efficient in promoting behavior change [14]. Yet it has not been largely assessed in the context of medication adherence. Moreover, to our knowledge, there are currently no apps with gamification that target the Swiss market with the available medications in this country [15,16].

In an attempt to boost adherence, a multidisciplinary team of health professionals, informaticians, and patients in a cardiac rehabilitation (CR) program worked together to develop an innovative app with gamification strategies named "Smart-Meds."

The main objective of this study was to evaluate the adoption, usability, and satisfaction of Smart-Meds among users enrolled in an outpatient CR program. We also explored the impact of app use on medication adherence and beliefs.

## Methods

### Study Design

This is a pilot pre-post study aimed at assessing the impact on participants' self-efficacy regarding their medication regimens and their beliefs about medication efficacy following the use of the Smart-Meds app for 1 month.

### Primary and Secondary Outcome

The primary outcome is the Self-Efficacy for Appropriate Medication Use Scale (SEAMS), and the secondary outcomes are the Beliefs About Medication Questionnaire (BMQ) and the System Usability Scale (SUS).

### Participants

We included adults (>18 years) who were treated for an ACS in the past month and who owned an Android or iPhone. We excluded participants who did not speak conversational French.

### Sample Size

In this pilot pre-post study, the sample size was determined using the rule of thumb for pilot studies, which suggests a minimum of 12 participants per group to provide an initial estimate of effect sizes and variability [17]. This sample size is considered adequate for assessing feasibility and refining study protocols, while not intended for definitive hypothesis testing. The selected sample size allows for the identification of trends and potential issues that may inform the design of a subsequent, fully powered study.

### Recruitment

We enrolled voluntary participants entering a CR program at the University Hospital of Geneva. Patients were recruited during round table sessions by an investigator presenting the study. After providing their consent, the participants received help if needed to install and use the app on their smartphones.

## Ethical Consideration

An ethical application was made to the hospital's ethical committee. The ethics committee considered that this research was targeted mainly to evaluate the application itself and could be considered as quality-related research. Therefore, they exempted us from ethical approval. Informed consent was signed by all participants prior to the inclusion in the study. All data collected in the study have been anonymized by using unique identifiers before analysis, ensuring that no personal information could be traced back to any individual. There was no need for compensation, and no images of individual participants were included in this paper and supplementary materials.

## Intervention

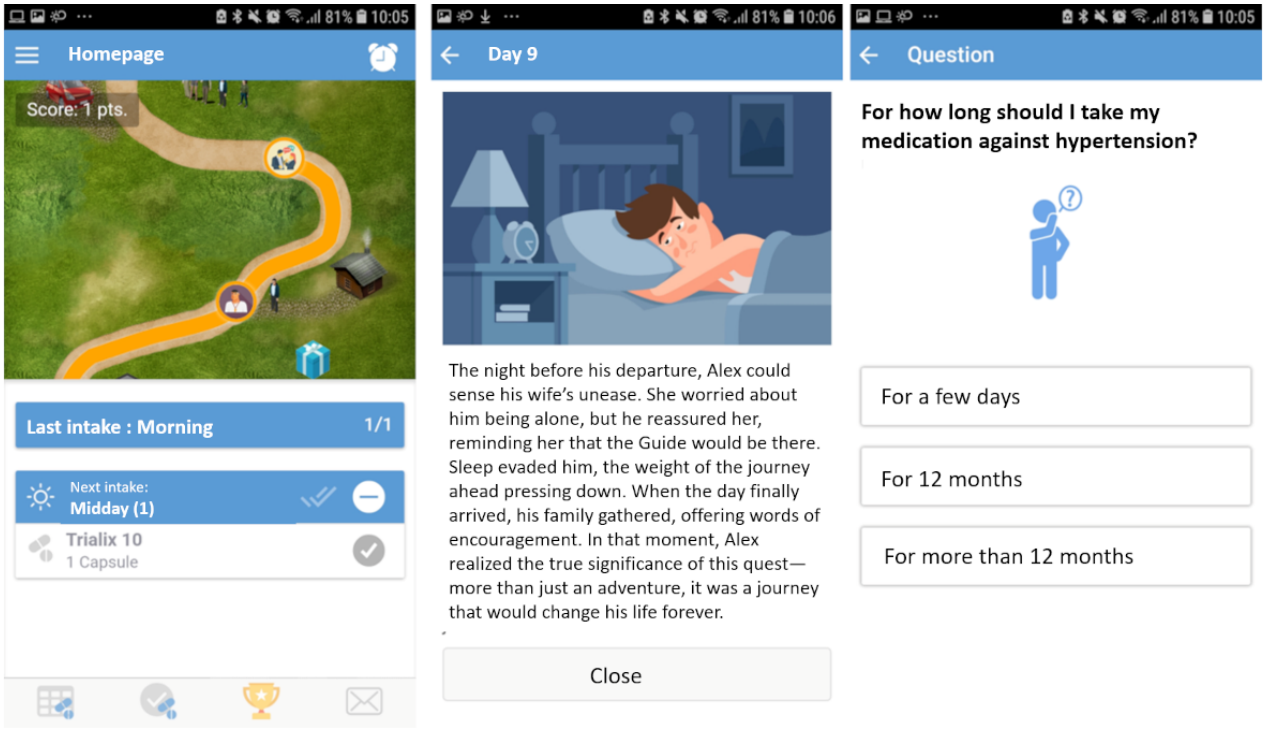
Smart-Meds is an app created following a participatory design. Users were involved all along its development, providing feedback at each step of the iterative cycles of formative evaluation [18]. The users participating in the app conception were patients participating in or having recently completed the 6-week CR program. The app's main aim is to empower users to manage their medications, using gamification strategies to motivate users to report their intakes. The app allows users to easily enter medications into their personal medication plan through barcode scanning of the drug boxes. Besides avoiding transcription errors, this process ensures that the correct medication is entered (pharmacies may provide different generics of a drug), and the user only has the dosage and schedule to enter. Users can set reminders about when to take their medications and have links to the Swiss patient information web page about their drugs. For the standard cardiovascular drugs, our team also developed simplified information content about indications and side effects that were adapted to low health literacy levels. We also created an educational section in the app about coronary heart disease, based on the CR program materials.

To increase users' motivation to report their medication intake, we relied on gamification mechanisms. The core mechanism is a narration whose daily stages of a motivational story are unlocked by reporting medication intake. Narrative has been demonstrated to be a relevant mechanism that can foster behavior change [19]. Narratives can help bridge the gap between intention and action. The health action process approach suggests that people may not act on a desired behavior for different reasons: those who are not (yet) motivated to do so are nonintenders, while intenders may be motivated but unable to put their intention into action [20]. According to this approach, planning strategies are essential in aiding intenders to close this gap. These strategies involve specifying when, where, and how to carry out the desired behavior (action planning) and anticipating potential obstacles and preparing ways to overcome them (coping planning). Narratives are particularly useful in this regard; they focus on specific characters, their actions and motivations, and present events in a temporal and causal structure. Therefore, characters can act as role models, demonstrating how to turn intention into action, what to expect in terms of challenges, and how to navigate them successfully [21].

This story was designed to increase engagement and reinforce the concepts of the "health action process approach" model [22] and is inspired by an annual outing for patients with ACS at the Cardiac Rehabilitation Center of the University Hospital of Geneva. The story consists of 30 episodes. The average textual length of each episode is 470 characters.

Another gamified mechanism implemented in the app is the progression since the user sees its progression toward storing through a visual path on the app (Figure 1).

**Figure 1.** Screenshot of the app: the first screen displays the story stages unlocked by reporting medication, the second screen displays a part of the story, and the third screen displays an example of the quiz (translated into English from the original French version).




Users can also test their knowledge about coronary heart disease and its management through daily quizzes. Finally, the app allows users to evaluate their cardiovascular risk factors to guide their lifestyle changes. A more detailed description of the app and its underlying framework is reported elsewhere [23].

### Study

#### *Measures and Data Collection*

Once recruited, participants completed questionnaires on demographic data and on medication adherence and beliefs (SEAMS and BMQ) [24,25]. SEAMS is a self-reported questionnaire with 13 items about how to manage one's drugs in various situations (eg, change in routine, suspected side effects, and new prescriptions). The BMQ has 18 items, with subsets of questions on the nature of medication, their use by doctors, one's personal need for a drug, and concerns about side effects. The participant then received the mobile app and received some help if necessary to install the app on their smartphones. The investigators also helped the participants to enter their treatment into the app. The participants were then instructed to use the app for 4 weeks at home without any interactions with the investigators or any recall.

After 4 weeks, in addition to the completion of a second SEAMS and BMQ, participants scored the app with the SUS. An investigator also conducted a semistructured oral interview in person or by phone. Nine open-ended questions were designed by the investigators based on a combination of deductive and inductive approaches. The investigation team started with the research objectives (deductive) and refined and expanded questions based on insights gained from initial data analysis and literature review (inductive). The selected questions explored reasons for satisfaction and app use and enquired about suggestions for improvements. The investigator audio-recorded the interviews or took session notes for a subsequent analysis. We also collected data about app use from the app logs (number of sessions, duration of session). Due to technical limitations, the log data were only captured when the participant was online at the time of app use. Only log sessions lasting more than 1 second were considered significant for this study.

#### *Data Analysis*

We report descriptive statistics of the demographic data to characterize our sample and of the use logs. We used a qualitative approach for the interviews, extracting common themes through iterative coding and comparisons of the data. SEAMS and BMQ scores are reported before and after the intervention and their distribution is compared using a chi-square analysis. Analyses were done using Microsoft Excel version 1808.

The study was carried out in French: as there was no validated translation available at the time of the study for the SEAMS, we proceeded with a translation or back-translation with 2 external consultants.

## Results

### Demographics

We recruited participants between February and April 2020. We report the results of the 18 participants who completed the study in Table 1 (of 37 participants screened for eligibility, 19 declined). Overall, participants were mainly male and Caucasian, with high socioeconomic status, which is representative of our targeted population. All participants had 4G connectivity. At the beginning of the study, half the participants monitored their blood pressure and physical activity.

**Table .** Participant characteristics (n=18).

| Variable | | Values |
|---|---|---|
| Week of program at enrollment (total of 6 weeks), mean (IQR) | | 2 (1-2.75) |
| Medications, mean (IQR) | | 5 (4.25-7.75) |
| **Age category (years), n (%)** | | |
| | 35-44 | 2 (11) |
| | 45-54 | 5 (28) |
| | 55-64 | 8 (44) |
| | 65-74 | 3 (17) |
| **Sex, n (%)** | | |
| | Male | 16 (89) |
| | Female | 2 (11) |
| **Educational attainment, n (%)** | | |
| | High school | 7 (39) |
| | College or higher | 11 (61) |
| **Origin, n (%)** | | |
| | Caucasian | 14 (78) |
| | Other | 4 (22) |
| **Private health insurance, n (%)** | | |
| | Yes | 12 (67) |
| | No | 6 (33) |
| **Type of smartphone, n (%)** | | |
| | Android | 7 (39) |
| | iPhone | 11 (61) |
| **Use of apps for health, n (%)** | | |
| | Wellness | 2 (11) |
| | Medical | 6 (33) |
| | None | 10 (56) |
| **Current monitored parameter, n (%)** | | |
| | Blood pressure | 10 (56) |
| | Weight | 7 (39) |
| | Physical activity | 9 (50) |
| | Diet | 6 (33) |
| | Blood glucose | 2 (11) |

## Usage Pattern

All 18 participants installed and used Smart-Meds successfully. We see in Figure 1 that although every participant installed the app on the first day, we had an immediate dropout of one-third of the users. After that, the use remains stable until day 25.

On average, active participants used the app 3.76 (SD 1.28) sessions per day with a total of 64.39 (SD 21.55) seconds per day (Table 2). The highest app use was on the first day with an average of 4.67 sessions per participant of 2.5 minutes duration. App use drops rapidly after the first couple of days and persists at about 1x/day until the end of the 30 days.

**Table .** Use of the Smart-Meds app over the 30 days.

| Day of the study | Daily user, n | Still active participants, n (%) | Sessions per active user, mean (SD) | App use duration per user (second), mean (SD) |
|---|---|---|---|---|
| 1 | 18 | 18 (100) | 4.67 (5.69) | 147.98 (267.91) |
| 2 | 11 | 13 (72) | 6.29 (5.55) | 89.31 (142.29) |
| 3 | 8 | 12 (67) | 3.13 (2.10) | 60.68 (90.08) |
| 4 | 9 | 12 (67) | 3.89 (2.71) | 67.96 (98.49) |
| 5 | 9 | 12 (67) | 4.71 (3.77) | 74.88 (130.38) |
| 6 | 9 | 12 (67) | 3.00 (2.00) | 85.26 (108.58) |
| 7 | 6 | 12 (67) | 4.89 (3.98) | 56.02 (55.53) |
| 8 | 9 | 12 (67) | 3.22 (2.33) | 57.05 (52.51) |
| 9 | 8 | 12 (67) | 2.29 (2.63) | 40.92 (44.42) |
| 10 | 8 | 12 (67) | 4.00 (3.34) | 80.94 (84.45) |
| 11 | 6 | 12 (67) | 3.00 (1.79) | 68.02 (61.53) |
| 12 | 9 | 12 (67) | 4.57 (1.72) | 42.53 (39.62) |
| 13 | 9 | 12 (67) | 5.50 (6.87) | 54.11 (60.09) |
| 14 | 8 | 12 (67) | 5.43 (6.24) | 43.88 (44.30) |
| 15 | 9 | 12 (67) | 4.43 (3.95) | 48.54 (60.97) |
| 16 | 9 | 12 (67) | 2.20 (1.99) | 66.43 (80.70) |
| 17 | 8 | 12 (67) | 3.00 (2.00) | 72.16 (62.64) |
| 18 | 8 | 12 (67) | 6.88 (7.49) | 48.56 (42.78) |
| 19 | 8 | 12 (67) | 3.71 (1.80) | 73.35 (96.89) |
| 20 | 9 | 12 (67) | 4.50 (4.47) | 75.48 (110.52) |
| 21 | 9 | 12 (67) | 2.00 (1.41) | 41.47 (33.05) |
| 22 | 6 | 12 (67) | 2.33 (1.53) | 77.99 (55.05) |
| 23 | 8 | 12 (67) | 2.20 (1.64) | 48.77 (35.50) |
| 24 | 7 | 12 (67) | 4.00 (3.70) | 60.96 (65.65) |
| 25 | 9 | 12 (67) | 3.50 (2.26) | 38.39 (50.23) |
| 26 | 10 | 12 (67) | 2.11 (1.17) | 77.14 (85.99) |
| 27 | 10 | 11 (61) | 2.43 (1.40) | 52.93 (43.41) |
| 28 | 8 | 9 (50) | 4.63 (3.66) | 76.94 (163.36) |
| 29 | 8 | 9 (50) | 3.71 (2.63) | 43.39 (48.26) |
| 30 | 7 | 7 (39) | 2.67 (1.86) | 59.55 (47.32) |

## Pre-Post Evaluation of SEAMS and BMQ

Although we did not find a significant change in the assessments of medical beliefs (BMQ, *P*=.09), the self-reported medication adherence score was significantly higher after 4 weeks (SEAMS, *P*=.02). Distribution of the SEAMS and BMQ scores can be visualized in Figures 2 and 3.

**Figure 2.** Boxplot of the BMQ score before and after the intervention period for the 18 participants. BMQ: Beliefs About Medication Questionnaire.

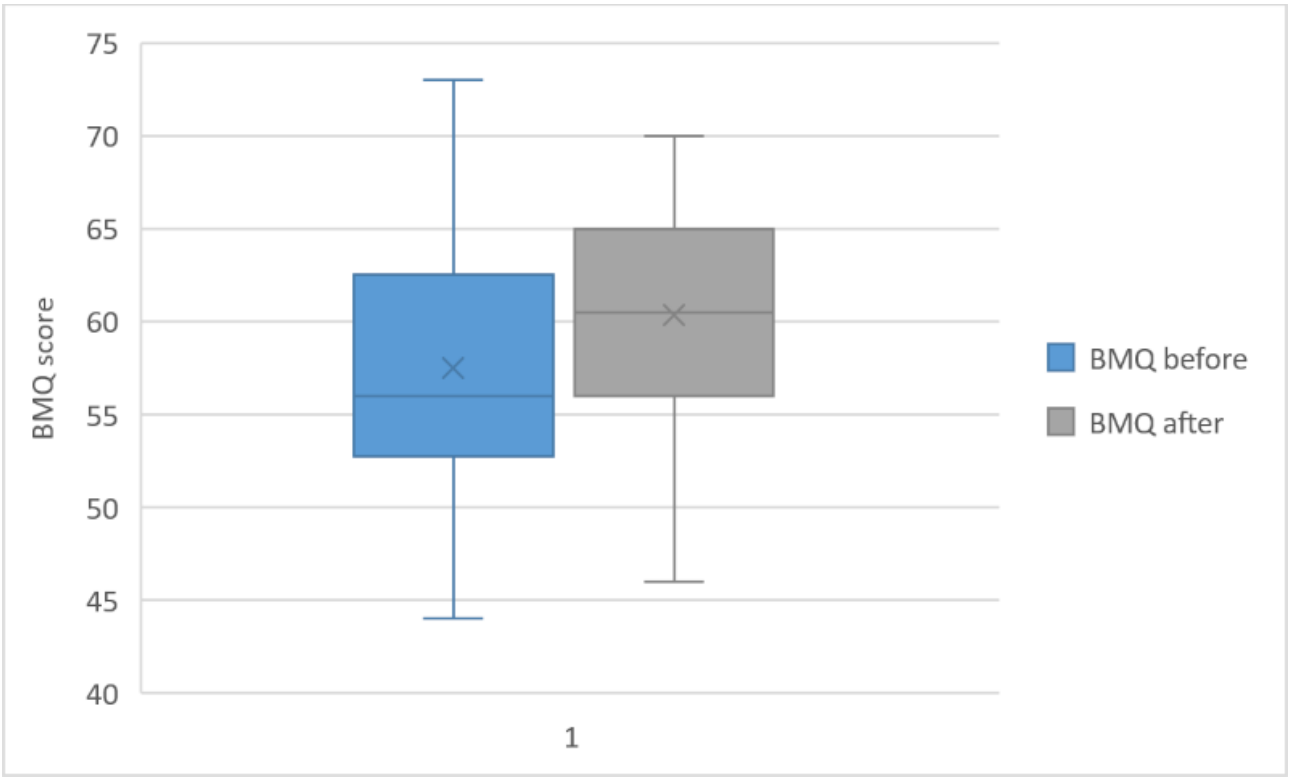
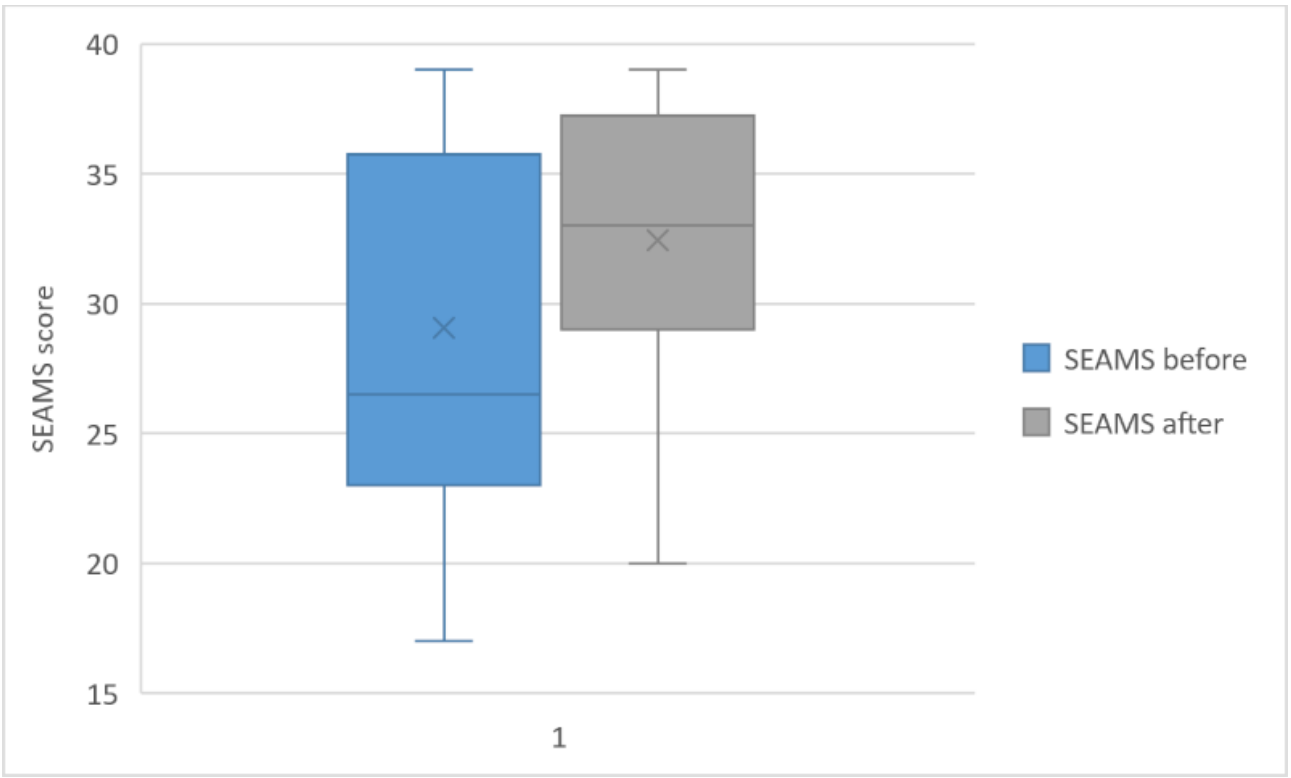



**Figure 3.** Boxplot of the SEAMS score before and after the intervention period for the 18 participants. SEAMS: Self-Efficacy for Appropriate Medication Use Scale.




## Semistructured Interview

In the semistructured interview, the 18 participants were overall very positive about the app, particularly when starting a new medication. Of the 18 participants, 5 (28%) liked being able to track their medication intake. One participant explained: “It’s very useful, because sometimes you can’t remember if you’ve taken the medication or not. With the app, I can validate taking the medication, and I do it as first action in the morning.” They were satisfied with the drug information and liked having an overview of all their medications, which they could share with their primary care physician. They appreciated its ease of use

and found the barcode scanning an easy and fun way to enter their medications in the app. Despite some bugs linked to the modification of the recall time in the reminder functionalities during the study, the users thought having reminders was useful. They also found having pictures of their medications useful, especially with new drugs.

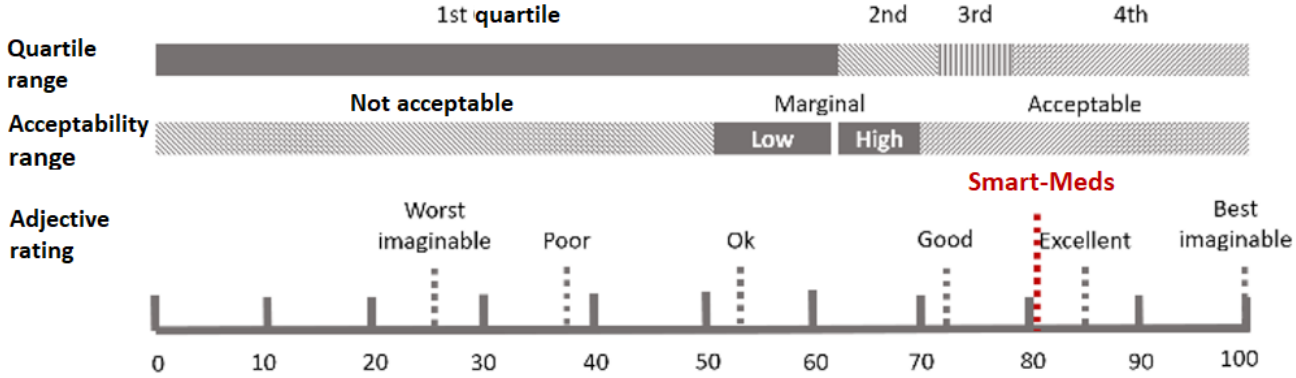
Of the 18 participants, 17 (94%) tested the quizzes and 15 (83%) enjoyed challenging their knowledge about their disease and their medications in this manner. In fact, 1 participant even suggested adding a reminder to take the quiz. Opinions about the motivational story were more varied because many participants did not engage with the story. Of the 18 participants, only 4 participants read the story until the end, and 1 participant suggested making it more interactive, where user choices affect the storyline. Half of the participants (9/18, 50%) reported the story as one of the less useful aspects of the app for them.

The participants did recognize that having a medication app was mainly useful early in the self-management process. Once they got into a routine to take the medication, the reminders were not as useful. In fact, 1 participant explained that taking his medications regularly was easy, but remembering to use the app was more difficult for him!

### System Usability Scale

Overall, the app was rated with a mean average score of 80.6 (SD 14.5), which may be interpreted as a good score according to Bangor et al [26]. The app was perceived between good and excellent (Figure 4).

**Figure 4.** Results of the System Usability Scale (SUS) for the 18 participants.




## Discussion

### Principal Findings

Our pilot study revealed that participant satisfaction among users was high and that they would recommend the app to others. Our results show an improvement in the self-reported medication adherence scale after 4 weeks of app use. Even though gamification has been demonstrated successful in boosting behavior change in several contexts, it seems to have a limited impact on our specific population.

### Comparison to Prior Work

Although several recent studies have suggested that gamification can drive health behavior change, the type of gamification technique needs to be considered [27,28]. For our participants, the impact of the motivational story was very different from the quiz. Storytelling was considered as a game, whereas the quiz was more a verification of acquired knowledge, something that they valued.

The story was created with ups and downs to represent daily variations when coping with a challenge. We kept the story sequences short and used many illustrations to draw the reader's attention. The users in our study did not demonstrate a strong interest in the motivational story. A plausible explanation is that the patients in our study were currently being treated for ACS, diagnosed in the past month [6]. We can suppose these participants were concerned about their current situation and did not find any added value from storytelling since their intrinsic motivation was already high [29,30].

The narrative approach has been used in other research. An article by Day [31] describes how storytelling has the potential to promote health literacy in patients. In the cardiology domain, Li et al [32] displayed an interactive video that depicted a model patient enacting a scenario with the patient experiencing acute myocardial infarction symptoms and going through the perceptual cognitive processes in decision-making. The psychoeducational intervention group reported greater positive changes than the control group in their attitudes.

The use of a quiz, however, another gamification technique, was well appreciated by the participants. Throughout the CR program, there are group discussions about heart disease, medications and side effects, and a healthy diet. They liked the idea of "checking" what knowledge they had acquired during the program. In fact, the quizzes were a way to monitor what they had understood and learned, rather than an outcome with the quiz score. Therefore, the participants had a much bigger interest in the quiz.

### Dropout

We observe in the log that one-third of the participants did only use the app once at the installation. This information does not correspond to the feedback of the patient during the semistructured interview. Indeed, during the interview, 14 patients reported using the app at least once per day, 3 patients twice per day, and 1 patient once every 2 days. The difference between the measured use and the reported one can have two reasons. First, research in various settings has demonstrated a difference between reported adherence and measured one [33]. The second reason is technical. Since the measure of adherence

is recorded on the backend, if the patient is not connected to the internet when reporting his or her intake, that information is not logged.

### Adherence

We observe that self-reported adherence to medication improved over time. Prior studies have shown that a good understanding of one's medication (why it is needed, how to take it, and potential side effects) is a driver for adherence [9,34]. Reading the simplified information facts in the app or self-testing with the quiz could have helped gain or maintain knowledge about medication during the study. Interestingly, the participants reported that the tracking functions were often not needed at this stage of their disease management: either they had already established a routine that suited them, or else they sometimes were low-tech and did not consider logging into the app regularly to track their medication intake [35,36]. Several participants considered this tracking as an additional, tedious task and therefore did not find tracking or reminders useful. The reminders were considered more useful when their routine was disrupted: this is commonly found in studies about adherence [37]. At this stage of the disease (CR program or right after the program), participants are still on sick leave at home, without the unexpected events that may occur from work-related tasks or travel issues.

### Other Contextual Elements

Participants enrolled in our study were from the CR program, with social support between peers, group sessions with health professionals, and daily physical activities in groups. In fact, patients often join a WhatsApp group to communicate with peers. This suggests other approaches to explore to help drive behavior changes, especially when the CR program ends, and "real life" begins again with work.

### Limitations

The first limitation of our study concerns the absence of a control group preventing to establish causality definitively. Without a control group for comparison, it becomes challenging to discern whether the observed changes in adherence behaviors and beliefs are solely attributable to the intervention or if they could be influenced by external factors or natural fluctuations over time. Additionally, the absence of a control group limits the researchers' ability to account for potential confounding variables that may impact the outcomes of interest. Therefore, while the pre-post pilot study design provides valuable insights into the potential effects of the intervention, its findings must be interpreted cautiously, and further research using a controlled study design is warranted to confirm and generalize the observed results.

The second limitation of this pre-post scientific pilot study is the small sample size, which may render the study underpowered. With a limited number of participants, the study's ability to detect significant changes in adherence behaviors and beliefs may be compromised. Small sample sizes can increase the likelihood of type II errors, where the study fails to detect real effects due to insufficient statistical power. Additionally, the generalizability of findings from a small sample size may be limited, as the characteristics and responses of a small group may not be representative of the broader population. Consequently, a cautious interpretation of the results is necessary, recognizing the potential limitations imposed by the small sample size on the study's reliability and generalizability. Future research with larger sample sizes would be beneficial to confirm and extend the findings of this pilot study.

Third, we faced limitations to record app use when offline. This may have led to a bias in the reporting of the results, as several users were voluntarily disconnecting their smartphones from wireless networks to minimize connection costs. Therefore, we can expect that users were using the app more frequently than reported.

### Future Direction

Building on the findings of this pilot study, future research could explore more tailored storytelling approaches to enhance patient engagement and adherence to medication. Identifying narratives that resonate more deeply with different patient populations may further improve the effectiveness of the gamified approach. Additionally, other gamification strategies, such as reward systems or adaptive challenges, could be investigated to assess their potential impact on patient outcomes.

A key next step is to conduct a larger-scale study with a control group to better assess the effectiveness of the gamified approach compared to traditional methods. This would allow for a more robust statistical analysis and provide stronger evidence of the intervention's benefits in improving medication adherence and patient awareness. Expanding the study to diverse patient demographics would also offer insights into the approach's generalizability and scalability.

### Conclusion

Smart-Meds is a promising app; although one-third of the participants dropped out immediately, the remaining participants used the app regularly. The satisfaction of users was high, and participants would recommend the app to others. Our results show an improvement in the self-reported medication adherence scale after 4 weeks of app use. Although gamification has been successful in boosting behavior change in several contexts, it seems to have a limited impact on our specific population. Therefore, additional research should be conducted with the end user to design a story that boosts their motivation. On the experimental side, a larger study with a controlled design like a randomized controlled trial is needed to confirm our results.

## Acknowledgments

We would like to extend our sincere gratitude to the nursing staff for their invaluable assistance in the recruitment process.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

Conceptualization: FE, KB
Methodology: FE, KB
Software: HH
Formal analysis: HH, FE, KB
Writing – original draft: FE
Writing – review & editing: FE, KB, LG, PM
Supervision: PM

## Conflicts of Interest

None declared.

## References


1. Rosen OZ, Fridman R, Rosen BT, Shane R, Pevnick JM. Medication adherence as a predictor of 30-day hospital readmissions. Patient Prefer Adherence 2017;11:801-810. [doi: 10.2147/PPA.S125672] [Medline: 28461742]
2. Hugtenburg JG, Timmers L, Elders PJ, Vervloet M, van Dijk L. Definitions, variants, and causes of nonadherence with medication: a challenge for tailored interventions. Patient Prefer Adherence 2013;7:675-682. [doi: 10.2147/PPA.S29549] [Medline: 23874088]
3. Aggarwal B, Pender A, Mosca L, Mochari-Greenberger H. Factors associated with medication adherence among heart failure patients and their caregivers. J Nurs Educ Pract 2015;5(3):22-27. [doi: 10.5430/jnep.v5n3p22] [Medline: 25635204]
4. Pantuzza LL, Ceccato MDGB, Silveira MR, Junqueira LMR, Reis AMM. Association between medication regimen complexity and pharmacotherapy adherence: a systematic review. Eur J Clin Pharmacol 2017 Nov;73(11):1475-1489. [doi: 10.1007/s00228-017-2315-2] [Medline: 28779460]
5. Chow CK, Brieger D, Ryan M, Kangaharan N, Hyun KK, Briffa T. Secondary prevention therapies in acute coronary syndrome and relation to outcomes: observational study. Heart Asia 2019;11(1):e011122. [doi: 10.1136/heartasia-2018-011122] [Medline: 30728864]
6. Bots SH, Inia JA, Peters SAE. Medication adherence after acute coronary syndrome in women compared with men: a systematic review and meta-analysis. Front Glob Womens Health 2021;2:637398. [doi: 10.3389/fgwh.2021.637398] [Medline: 34816194]
7. Le RJ, Cullen MW, Lahr BD, Wright RS, Kopecky SL. Side effects of CV medications following hospitalization for ACS are associated with more frequent health-care contacts. J Cardiovasc Pharmacol Ther 2017 May;22(3):250-255. [doi: 10.1177/1074248416672009] [Medline: 27698079]
8. Bassand JP. Drug interactions in the setting of acute coronary syndromes and dual anti-platelet therapy. Eur Heart J Suppl 2006 Oct 1;8:G35-G37. [doi: 10.1093/eurheartj/sul053]
9. Kvarnström K, Westerholm A, Airaksinen M, Liira H. Factors contributing to medication adherence in patients with a chronic condition: a scoping review of qualitative research. Pharmaceutics 2021 Jul 20;13(7):1100. [doi: 10.3390/pharmaceutics13071100] [Medline: 34371791]
10. Armitage LC, Kassavou A, Sutton S. Do mobile device apps designed to support medication adherence demonstrate efficacy? A systematic review of randomised controlled trials, with meta-analysis. BMJ Open 2020 Jan 30;10(1):e032045. [doi: 10.1136/bmjopen-2019-032045] [Medline: 32005778]
11. Thakkar J, Kurup R, Laba TL, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
12. de Waure C, Lauret GJ, Ricciardi W, et al. Lifestyle interventions in patients with coronary heart disease: a systematic review. Am J Prev Med 2013 Aug;45(2):207-216. [doi: 10.1016/j.amepre.2013.03.020] [Medline: 23867029]
13. Al-Rayes S, Al Yaqoub FA, Alfayez A, et al. Gaming elements, applications, and challenges of gamification in healthcare. Inform Med Unlocked 2022;31:100974. [doi: 10.1016/j.imu.2022.100974]
14. Bassanelli S, Vasta N, Bucchiarone A, Marconi A. Gamification for behavior change: a scientometric review. Acta Psychol (Amst) 2022 Aug;228:103657. [doi: 10.1016/j.actpsy.2022.103657] [Medline: 35767927]
15. Santo K, Chow CK, Thiagalingam A, Rogers K, Chalmers J, Redfern J. MEDication reminder APPs to improve medication adherence in Coronary Heart Disease (MedApp-CHD) Study: a randomised controlled trial protocol. BMJ Open 2017 Oct 8;7(10):e017540. [doi: 10.1136/bmjopen-2017-017540] [Medline: 28993388]
16. Burn E, Nghiem S, Jan S, et al. Cost-effectiveness of a text message programme for the prevention of recurrent cardiovascular events. Heart 2017 Jun;103(12):893-894. [doi: 10.1136/heartjnl-2016-310195] [Medline: 28235776]
17. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]

18. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579. [doi: 10.2196/11579] [Medline: 30552080]
19. Hinyard LJ, Kreuter MW. Using narrative communication as a tool for health behavior change: a conceptual, theoretical, and empirical overview. Health Educ Behav 2007 Oct;34(5):777-792. [doi: 10.1177/1090198106291963]
20. Schwarzer R, Hamilton K. Changing behavior using the health action process approach. In: The Handbook of Behavior Change: Cambridge University Press; 2020:89-103. [doi: 10.1017/9781108677318.007]
21. Boeijinga A, Hoeken H, Sanders J. Storybridging: four steps for constructing effective health narratives. Health Educ J 2017 Dec;76(8):923-935. [doi: 10.1177/0017896917725360] [Medline: 29276232]
22. Zhang CQ, Zhang R, Schwarzer R, Hagger MS. A meta-analysis of the health action process approach. Health Psychol 2019;38(7):623-637. [doi: 10.1037/hea0000728]
23. Ehrler F, Gschwind L, Meyer P, Christian L, Blondon K. SMART-MEDS: development of a medication adherence app for acute coronary syndrome patients based on a gamified behaviour change model. AMIA Annu Symp Proc 2018;2018:413-421. [Medline: 30815081]
24. Risser J, Jacobson TA, Kripalani S. Development and psychometric evaluation of the Self-Efficacy for Appropriate Medication Use Scale (SEAMS) in low-literacy patients with chronic disease. J Nurs Meas 2007;15(3):203-219. [doi: 10.1891/106137407783095757] [Medline: 18232619]
25. Allen LaPointe NM, Ou FS, Calvert SB, et al. Association between patient beliefs and medication adherence following hospitalization for acute coronary syndrome. Am Heart J 2011 May;161(5):855-863. [doi: 10.1016/j.ahj.2011.02.009] [Medline: 21570514]
26. Bangor A, Staff T, Kortum P, et al. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123. [doi: 10.5555/2835587.2835589]
27. Edwards EA, Lumsden J, Rivas C, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. BMJ Open 2016 Oct 4;6(10):e012447. [doi: 10.1136/bmjopen-2016-012447] [Medline: 27707829]
28. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3. [doi: 10.2196/games.3139] [Medline: 25658754]
29. Greco A, Cappelletti ER, Monzani D, et al. A longitudinal study on the information needs and preferences of patients after an acute coronary syndrome. BMC Fam Pract 2016 Sep 20;17:136. [doi: 10.1186/s12875-016-0534-8] [Medline: 27646507]
30. Kähkönen O, Kankkunen P, Saaranen T, Miettinen H, Kyngäs H, Lamidi ML. Motivation is a crucial factor for adherence to a healthy lifestyle among people with coronary heart disease after percutaneous coronary intervention. J Adv Nurs 2015 Oct;71(10):2364-2373. [doi: 10.1111/jan.12708] [Medline: 26084708]
31. Day V. Promoting health literacy through storytelling. Online J Issues Nurs 2009;14(3). [doi: 10.3912/OJIN.Vol14No03Man06]
32. Li PWC, Yu DSF, Yan BP, et al. Effects of a narrative-based psychoeducational intervention to prepare patients for responding to acute myocardial infarction: a randomized clinical trial. JAMA Netw Open 2022;5(10):e2239208. [doi: 10.1001/jamanetworkopen.2022.39208]
33. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. J Med Internet Res 2022 Apr 26;24(4):e35120. [doi: 10.2196/35120] [Medline: 35471414]
34. Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. Patient Educ Couns 2016 Jul;99(7):1079-1086. [doi: 10.1016/j.pec.2016.01.020] [Medline: 26899632]
35. Gandapur Y, Kianoush S, Kelli HM, et al. The role of mHealth for improving medication adherence in patients with cardiovascular disease: a systematic review. Eur Heart J Qual Care Clin Outcomes 2016 Oct 1;2(4):237-244. [doi: 10.1093/ehjqcco/qcw018] [Medline: 29474713]
36. van Heuckelum M, van den Ende CHM, Houterman AEJ, Heemskerk CPM, van Dulmen S, van den Bemt BJF. The effect of electronic monitoring feedback on medication adherence and clinical outcomes: a systematic review. PLoS One 2017;12(10):e0185453. [doi: 10.1371/journal.pone.0185453] [Medline: 28991903]
37. Choudhry NK, Krumme AA, Ercole PM, et al. Effect of reminder devices on medication adherence: the REMIND randomized clinical trial. JAMA Intern Med 2017 May 1;177(5):624-631. [doi: 10.1001/jamainternmed.2016.9627] [Medline: 28241271]


## Abbreviation

**ACS:** acute coronary syndrome
**BMQ:** Beliefs About Medication Questionnaire
**CR:** cardiac rehabilitation
**SEAMS:** Self-Efficacy for Appropriate Medication Use Scale
**SUS:** System Usability Scale

*Edited by A Coristine; submitted 10.07.23; peer-reviewed by C Eaton, M Cozad; revised version received 07.10.24; accepted 07.10.24; published 23.01.25.*

*Please cite as:*
*Ehrler F, Gschwind L, Hagberg H, Meyer P, Blondon K*
*A Medication Management App (Smart-Meds) for Patients After an Acute Coronary Syndrome: Pilot Pre-Post Mixed Methods Study*
*JMIR Cardio 2025;9:e50693*
*URL: https://cardio.jmir.org/2025/1/e50693*
*doi:10.2196/50693*

# Validity of Heart Rate Measurement Using Wearable Devices During Cardiopulmonary Exercise Testing in Patients With Cardiovascular Disease: Prospective Pilot Validation Study

Kazufumi Kitagaki[1,2], PhD; Yuji Hongo[2], MSc; Rie Futai[3], PhD; Takeshi Hasegawa[2], RPT; Hiroshi Morikawa[2], BSc; Hisashi Shimoyama[3], PhD

[1]Faculty of Rehabilitation, Shijonawate Gakuen University, 5-11-10 Houjou Daito, Osaka, Japan

[2]Department of Medical Technology, Itami City Hospital, Itami, Japan

[3]Department of Cardiovascular Medicine, Itami City Hospital, Itami, Japan

**Corresponding Author:**
Kazufumi Kitagaki, PhD
Faculty of Rehabilitation, Shijonawate Gakuen University, 5-11-10 Houjou Daito, Osaka, Japan

## Abstract

**Background:** Wearable devices offer a promising solution for remotely monitoring heart rate (HR) during home-based cardiac rehabilitation. However, evidence regarding their accuracy across varying exercise intensities and patient profiles remains limited, particularly in populations with cardiovascular disease (CVD) such as those with heart failure (HF).

**Objective:** The objective of this study was to evaluate the accuracy of HR measurements obtained using the Fitbit Inspire 3 during cardiopulmonary exercise testing (CPX) in patients with CVD, including those with HF.

**Methods:** In this single-center, prospective pilot study, we enrolled 30 patients with CVD undergoing CPX. HR was simultaneously recorded using electrocardiography and the Fitbit Inspire 3 at 1-minute intervals across various CPX phases: rest, exercise below and above the anaerobic threshold (AT), and recovery. The correlation between the two methods was assessed using the Pearson correlation coefficient. Measurement error was quantified by mean absolute error and mean absolute percentage error (MAPE), with a MAPE of ≤10% defined as the threshold for acceptable agreement.

**Results:** All data points were 630 points per minute. The Fitbit Inspire 3 device demonstrated a strong overall correlation with electrocardiography-derived HR ($r$=0.90; IQR 0.88 - 0.91) and an acceptable MAPE of 5.40% (SD 8.33%). The total error was 14.9% (94/630), with overestimation and underestimation of 37 (5.8%) points and 57 (9%) points, respectively. The rate of HR underestimation reached 19 (16%) points during exercise above the AT, compared to 1 (3%) point at rest. When stratified by HF stage (B vs C), underestimation was more pronounced in patients with HF (14/275, 5% points vs 40/355, 11.2% points).

**Conclusions:** The Fitbit Inspire 3 provides acceptable validity for HR monitoring during CPX in patients with CVD. However, clinicians should interpret HR data with caution during high-intensity exercise, especially in patients with HF.

***(JMIR Cardio 2025;9:e77911)*** doi:10.2196/77911

**KEYWORDS**

validation; heart rate measurement; wearable device; Fitbit Inspire 3; heart failure

## Introduction

### Background

Outpatient cardiac rehabilitation (CR) is a class I recommended therapy for patients with cardiovascular disease (CVD) and is regarded as an essential component of treatment [1]. However, in Japan, the prevalence of heart failure (HF) is rising among older adults [2], while participation in outpatient CR remains low [3]. Among working-age patients with coronary artery disease (CAD), the challenge of balancing work and CR contributes to the low participation rates [4]. A recent meta-analysis [5] demonstrated that home-based CR combined with digital support yielded comparable improvements in quality of life and reductions in hospital readmissions to conventional outpatient CR. These findings highlight the potential significance of remote support in home-based CR programs.

Maintaining appropriate exercise intensity is crucial to prevent symptom exacerbation when prescribing exercise therapy [1]. Heart rate (HR), closely linked to oxygen consumption [6], is a key indicator for setting exercise intensity and is commonly used in outpatient CR [7]. HR monitoring during outpatient CR is typically performed using electrocardiography (ECG). However, in home-based settings, continuous ECG monitoring is impractical, leading to reliance on manual pulse checks, whose accuracy remains insufficiently validated.

Advances in wearable digital technology have led to the widespread use of devices using photoplethysmography (PPG), such as smartwatches, which are being investigated as potential alternatives to ECG [8]. A recent scoping review suggested that smartwatches may be useful tools to support and enhance outpatient disease management [9]. However, measurement accuracy can vary widely between devices [10], and HR obtained via wearable devices tends to be underestimated at higher exercise intensities [11]. Consequently, whether such technology can be reliably applied to patients with CVD, particularly those with HF who may experience reduced peripheral perfusion, warrants careful evaluation.

A recent validation study [12] reported that HR measurements using the Apple Watch 7 (Apple Inc) and Galaxy Watch 4 (Samsung Electronics Co, Ltd) during cardiopulmonary exercise testing (CPX) were highly accurate in patients with CAD, indicating their potential utility in this population. However, these devices are relatively expensive (approximately US $350), and their accessibility to individuals with low income, who may already be reluctant to participate in outpatient CR, remains limited [4]. Moreover, although the study population included stable patients with CAD, the accuracy of measurement in patients with HF was not evaluated.

### Objective

We hypothesized that if a more affordable device, such as the Fitbit Inspire 3 (Fitbit LLC; approximately US $80), could yield similar HR measurement accuracy, it would become more widely accessible. Previous research on the Fitbit series in postoperative patients performing low-intensity activities [13] showed promising accuracy, and a validation study involving younger patients with CVD demonstrated good performance during stepwise increases in exercise intensity measured by CPX [14]. However, no studies have specifically focused on patients with HF. Therefore, this study aimed to evaluate the accuracy of HR measurement from the Fitbit Inspire 3 compared to ECG-based measurements during CPX in patients with CVD, including those with HF.

## Methods

### Study Design and Setting

This single-center, prospective, observational pilot study was conducted at the Itami City Hospital.

### Participants

Participants were recruited from patients who received a prescription for CPX at the Department of Cardiology, Itami City Hospital, between August 2024 and March 2025. The inclusion criteria were as follows: (1) aged ≥18 years at the time of consent, (2) documented agreement to undergo CPX, and (3) provision of written informed consent for study participation. The exclusion criteria were as follows: (1) aged <18 years, (2) persistent atrial fibrillation (AF), (3) known silicone allergy, or (4) refusal to participate.

### Sample Size Calculation

As a pilot study, a target sample size of 30 participants was established based on previous studies with sample sizes ranging from 10 to 60 [10], as well as the estimated number of CPX procedures expected during the study period.

### Data Collection and Procedures

CPX was performed using a cycle ergometer equipped with a breath-by-breath gas analyzer (AE-300S; Minato Medical Science). The testing protocol included a 1-minute rest period and 4 minutes of warm-up at 0 W, followed by a symptom-limited ramp protocol with individualized increments of 10 - 20 W/min. The recovery phase involved pedaling at 0 W for ≥5 minutes until clinical status stabilized. Ventilatory parameters, including minute ventilation (VE), oxygen uptake ($VO_2$), and carbon dioxide output ($VCO_2$), were recorded every 6 seconds.

The respiratory exchange ratio was calculated as $VCO_2/VO_2$ using breath-by-breath data. Peak $VO_2$ was defined as the highest $VO_2$ value recorded during exercise or the average of the final 18 seconds (3 data points), whichever was greater. Peak $VO_2$ was normalized by body weight (BW) and expressed in mL/kg/min (peak $VO_2$/BW), in accordance with standard clinical practice for CPX. The anaerobic threshold (AT) was determined by an experienced cardiologist based on multiple indices: the inflection point in $VE/VO_2$ without a concurrent rise in $VE/VCO_2$, increasing partial pressure of end-tidal oxygen without a change in end-tidal carbon dioxide, and other standard criteria.

During CPX, the Fitbit Inspire 3 device was attached to the arm contralateral to the side used for blood pressure measurement. The device was worn approximately 2 finger-widths above the wrist crease, with the optical sensor positioned flush against the skin to minimize motion artifacts. HR was recorded simultaneously using ECG and the wearable device, which logged data at 1-minute intervals [15]. Test start times were extracted from the electronic medical records. HR data from the wearable device were retrieved through the Fitbit Web Application Programming Interface using the *httr* and *jsonlite* packages in R (version 4.4.0; R Foundation for Statistical Computing).

### Demographic and Clinical Characteristics

All variables were obtained from patients' medical records. Baseline demographics data included age, sex, diagnosis, American College of Cardiology (ACC) or American Heart Association (AHA) HF stages [16], relevant medical history, current medications, and smoking status. Physical examination findings included height, weight, BMI, and the New York Heart Association functional classification. Laboratory parameters included hemoglobin, creatinine, estimated glomerular filtration rate, and B-type natriuretic peptide. Echocardiographic data included left ventricular ejection fraction (LVEF), measured using the modified Simpson method. HF phenotypes were categorized as HF with reduced ejection fraction (LVEF≤40%), HF with mildly reduced ejection fraction (LVEF 41% - 49%), and HF with preserved ejection fraction (LVEF≥50%) [17].

### Statistical Analysis

Continuous variables are presented as mean (SD) or median (IQR), as appropriate. Categorical variables are expressed as counts and percentages.

A whole-test analysis was performed using the complete set of 1-minute HR data. In addition, a phase-specific analysis was conducted by assigning each 1-minute sample to 1 of 4 predefined exercise phases: rest, low-intensity (including warm-up; below AT), high-intensity (above AT), and recovery. A stratified whole-test analysis was also carried out based on the ACC or AHA HF stage classification.

Furthermore, stratified analyses by ACC or AHA stage were conducted by repeating both the whole-test and phase-specific analyses within each stage category. Pearson correlation coefficients were calculated to assess the relationship between ECG-based and Fitbit-derived HR. Correlation coefficients were interpreted as follows: 0 - 0.30 as negligible, 0.30 - 0.50 as low, 0.50 - 0.70 as moderate, 0.70 - 0.90 as high, and 0.90 - 1.00 as very high. Bland-Altman plots were used to evaluate agreement, systematic bias, and limits of agreement. Mean absolute error (MAE) and mean absolute percentage error (MAPE) were calculated to quantify measurement error. On the basis of previous studies, a MAPE of ≤10% was considered acceptable [18,19]. The proportions of underestimation and overestimation were also calculated. All numerical values and statistical metrics were derived using CPX-measured HR as the reference standard. All analyses were conducted using R (version 4.4.0; R Foundation for Statistical Computing) and RStudio (version 2024.12.1+563; RStudio, Inc).

### Ethical Considerations

The study protocol complied with the principles of the Declaration of Helsinki (1975), as revised in 2000, and was approved by the institutional review boards of the Itami City Hospital (2537) and the Shijonawate Gakuen University (24 - 4). Written informed consent was obtained from all participants after detailed explanations of the study's objectives, procedures, potential benefits, and risks were provided by the researchers. To protect participant privacy, study records were de-identified. Each participant was assigned a unique study ID, and the linkage file connecting IDs to personal identifiers was stored exclusively on the hospital's electronic medical record network; no identifiable information was transferred outside this network. Only variables necessary for the study were collected, and the datasets were processed and managed to prevent the immediate identification of any individual. All analyses were conducted using de-identified datasets on the local computers. Participants did not receive any financial or material compensation for participating in the study.

## Results

### Demographic and Clinical Characteristics

The demographic and clinical characteristics are summarized in Tables 1 and 2. Of the 30 patients with CVD included in the study, 13 (43%) were classified as having stage B HF and 17 (57%) as having stage C HF. The median (IQR) age was 65 (54-73) years, and 17 patients (57%) were male. CAD was the most common underlying condition, observed in 16 patients (53%), followed by dilated cardiomyopathy in 6 (20%) patients. Regarding pharmacotherapy, β blockers were prescribed to 23 (77%) patients. According to the New York Heart Association classification, 8 (27%) patients were in class I; 20 (67%) in class II; and 2 (7%) in class III. The median (IQR) LVEF was 52% (45% - 61%), and 20% (6/30) of the patients were categorized as having HF with reduced ejection fraction. The median (IQR) B-type natriuretic peptide concentration was 56 (13 - 93) pg/mL. During CPX, the total median exercise time was 8.33 (IQR 7.00 - 10.45) minutes, the median peak respiratory exchange ratio was 1.16 (IQR 1.10 - 1.23), and the median peak $VO_2$ /BW was 17.2 (IQR 14.5 - 21.2) mL/kg/min. When comparing the stage B (n=13, 43%) and stage C (n=17, 57%) groups, patients in the stage B group were older (median age 73 vs 55 y) and had a higher prevalence of CAD (11/13, 85% vs 5/17, 29%). In contrast, nonischemic etiologies, including dilated cardiomyopathy, were more common in the stage C group (6/17, 35% vs 0/13, 0%). β-blocker use was more frequent in the stage C group (17/17, 100% vs 6/13, 46%).

**Table .** Participant demographics and clinical profile.

| Characteristics | Overall (n=30) | Stage B (n=13) | Stage C (n=17) |
|---|---|---|---|
| Age (y), median (IQR) | 65 (54-73) | 73 (64-78) | 55 (48-65) |
| Male, n (%) | 17 (57) | 8 (62) | 9 (53) |
| BMI (kg/m$^2$), median (IQR) | 23.2 (21.6-26.2) | 22.0 (20.8-24.8) | 23.7 (22.0-27.9) |
| Etiology, n (%) | | | |
| CAD[a] | 16 (53) | 11 (85) | 5 (29) |
| DCM[b] | 6 (20) | 0 (0) | 6 (35) |
| HHD[c] | 2 (7) | 0 (0) | 2 (12) |
| Cardiac sarcoidosis | 2 (7) | 0 (0) | 2 (12) |
| Others | 4 (13) | 2 (15) | 2 (12%) |
| Comorbidities, n (%) | | | |
| Hypertension | 19 (63) | 10 (77) | 9 (53) |
| Dyslipidemia | 12 (40) | 6 (46) | 6 (35) |
| Smoking, n (%) | | | |
| Never | 18 (60) | 8 (62) | 10 (59) |
| Past | 8 (27) | 4 (31) | 4 (24) |
| Current | 4 (13) | 1 (8) | 3 (18) |
| FH[d], n (%) | 1 (3) | 1 (8) | 0 (0) |
| DM[e], n (%) | 9 (30) | 4 (31) | 5 (29) |
| Respiratory disease, n (%) | 4 (13) | 2 (15) | 2 (12) |
| LVEF[f] (%), median (IQR) | 52 (45-61) | 54 (52-61) | 46 (39-58) |
| LVEF classification, n (%) | | | |
| HFpEF[g] | 16 (53) | 10 (77) | 6 (35) |
| HFmrEF[h] | 8 (27) | 3 (23) | 5 (29) |
| HFrEF[i] | 6 (20) | 0 (0) | 6 (35) |
| NYHA[j] classification, n (%) | | | |
| 1 | 8 (27) | 4 (31) | 4 (24) |
| 2 | 20 (67) | 9 (69) | 11 (65) |
| 3 | 2 (7) | 0 (0) | 2 (12) |

[a]CAD: coronary artery disease.
[b]DCM: dilated cardiomyopathy.
[c]HHD: hypertensive heart disease.
[d]FH: family history.
[e]DM: diabetes mellitus.
[f]LVEF: left ventricular ejection fraction.
[g]HFpEF: heart failure with preserved ejection fraction.
[h]HFmrEF: heart failure with mildly reduced ejection fraction.
[i]HFrEF: heart failure with reduced ejection fraction.
[j]NYHA: New York Heart Association.

**Table .** Pharmacotherapy, laboratory biomarkers, and CPX[a] parameters.

| Characteristics | Overall (n=30) | Stage B (n=13) | Stage C (n=17) |
|---|---|---|---|
| ACE-i[b], ARB[c], or ARNI[d], n (%) | 28 (93) | 13 (100) | 15 (88) |
| β blocker, n (%) | 23 (77) | 6 (46) | 17 (100) |
| SGLT2i[e], n (%) | 13 (43) | 2 (15) | 11 (65) |
| MRA[f], n (%) | 14 (47) | 1 (8) | 13 (76) |
| Diuretics, n (%) | 7 (23) | 0 (0) | 7 (41) |
| Nitrates, n (%) | 11 (37) | 8 (62) | 3 (18) |
| Calcium antagonists, n (%) | 3 (10) | 1 (8) | 2 (12) |
| Antiplatelet agents, n (%) | 16 (53) | 12 (92) | 4 (24) |
| Anticoagulants, n (%) | 8 (27) | 3 (23) | 5 (29) |
| Statins, n (%) | 16 (53) | 11 (85) | 5 (29) |
| Ezetimibe, n (%) | 1 (3) | 1 (8) | 0 (0) |
| Amiodarone, n (%) | 1 (3) | 0 (0) | 1 (6) |
| Hemoglobin (g/dL), median (IQR) | 13.75 (11.83-14.98) | 12.90 (11.80-14.80) | 13.80 (12.30-15.00) |
| Creatinine (mg/dL), median (IQR) | 0.92 (0.80-1.04) | 0.88 (0.80-1.11) | 0.92 (0.84-1.01) |
| eGFR[g] (mL/min/1.73 m²), median (IQR) | 62 (50-70) | 61 (50-69) | 64 (51-70) |
| BNP[h] (pg/mL), median (IQR) | 56 (13-93) | 61 (25-114) | 27 (12-83) |
| Missing BNP values, n (%) | 1 (3) | 1 (8) | 0 (0) |
| Ramp protocol | | | |
| 10 W, n (%) | 26 (87) | 12 (92) | 14 (82) |
| 20 W, n (%) | 4 (13) | 1 (8) | 3 (18) |
| Total exercise time (min), median (IQR) | 8.33 (7.00-10.45) | 8.90 (7.00-10.60) | 8.20 (7.15-9.65) |
| Peak RER[i], median (IQR) | 1.16 (1.10-1.23) | 1.16 (1.11-1.23) | 1.15 (1.04-1.22) |
| Peak WR[j] (W), median (IQR) | 93 (75-112) | 93 (81-110) | 98 (74-112) |
| Peak $VO_2$[k] (mL/min), median (IQR) | 1158 (883-1324) | 1097 (984-1,277) | 1194 (820-1348) |
| Peak $VO_2$/BW[l] (mL/kg/min), median (IQR) | 17.2 (14.5-21.2) | 17.1 (15.6-22.1) | 17.3 (14.4-20.2) |
| $ATVO_2$[m] (ml/min), median (IQR) | 724 (610-823) | 682 (624-772) | 745 (584-840) |
| $ATVO_2$/BW (mL/kg/min), median (IQR) | 11.80 (9.86-12.82) | 11.69 (10.64-12.54) | 11.99 (9.52-12.83) |

[a]CPX: cardiopulmonary exercise testing.
[b]ACE-i: angiotensin-converting enzyme inhibitor.
[c]ARB: angiotensin II receptor blocker.
[d]ARNI: angiotensin receptor–neprilysin inhibitor.
[e]SGLT2i: sodium-glucose cotransporter 2 inhibitor.
[f]MRA: mineralocorticoid receptor antagonist.
[g]eGFR: estimated glomerular filtration rate.
[h]BNP: B-type natriuretic peptide.
[i]RER: respiratory exchange ratio.
[j]WR: work rate.
[k]$VO_2$ : oxygen uptake.
[l]$VO_2$ /BW: oxygen uptake per body weight.
[m]$ATVO_2$ : oxygen consumption at anaerobic threshold.

### Validity of HR Measurement Using the Fitbit Inspire 3

A comparison between HR measurements obtained via CPX and those estimated by the Fitbit Inspire 3 is presented. In total, 630 data points were analyzed and categorized into 4 exercise phases: rest, below and above the AT, and recovery (Table 3). These data were further stratified by HF stage (stages B and C; Table 4).

**Table .** Accuracy of Fitbit HR[a] compared to CPX[b]-measured HR. A cutoff value of 10% for MAPE[c] was adopted from previous studies to evaluate the error rate.

| Condition | Correlation coefficient, *r* (95% CI) | MAE[d] (bpm[e]), mean (SD) | MAPE (%), mean (SD) | Error rate, n (%) | Overestimation, n (%) | Underestimation, n (%) |
|---|---|---|---|---|---|---|
| All (N=630)[f] | 0.90 (0.88 - 0.91) | 5.20 (8.75) | 5.40 (8.33) | 94 (15) | 37 (6) | 57 (9) |
| Rest (n=30)[f] | 0.92 (0.84 - 0.96) | 3.40 (4.15) | 4.67 (5.53) | 3 (10) | 2 (7) | 1 (3) |
| Below AT[g] (n=270)[f] | 0.85 (0.82 - 0.88) | 4.41 (7.28) | 5.23 (8.29) | 39 (14) | 10 (4) | 29 (11) |
| Above AT (n=119)[f] | 0.76 (0.67 - 0.82) | 8.03 (14.08) | 6.14 (9.62) | 20 (17) | 1 (1) | 19 (16) |
| Recovery (n=211)[f] | 0.92 (0.90 - 0.94) | 4.85 (6.46) | 5.32 (7.96) | 27 (13) | 22 (10) | 5 (2) |

[a]HR: heart rate.

[b]CPX: cardiopulmonary exercise testing.

[c]MAPE: mean absolute percentage error.

[d]MAE: mean absolute error.

[e]bpm: beats per minute.

[f]Number of data points.

[g]AT: anaerobic threshold.

**Table .** Accuracy of Fitbit HR[a] compared to CPX[b]-measured HR stratified by heart failure stage. A cutoff value of 10% for MAPE[c] was adopted from previous studies to evaluate the error rate.

| Condition | Correlation coefficient, *r* (95% CI) | MAE[d] (bpm[e]), mean (SD) | MAPE (%), mean (SD) | Error rate, n (%) | Overestimation, n (%) | Underestimation, n (%) |
|---|---|---|---|---|---|---|
| All (N=630)[f] | 0.90 (0.88 - 0.91) | 5.20 (8.75) | 5.40 (8.33) | 94 (15) | 37 (6) | 57 (9) |
| Stage B (n=275)[f] | 0.92 (0.90 - 0.94) | 4.51 (7.66) | 4.67 (7.94) | 31 (11) | 17 (6) | 14 (5) |
| Stage C (n=355)[f] | 0.88 (0.85 - 0.90) | 5.73 (9.49) | 5.98 (8.59) | 58 (16) | 18 (5) | 40 (11) |

[a]HR: heart rate.

[b]CPX: cardiopulmonary exercise testing.

[c]MAPE: mean absolute percentage error.

[d]MAE: mean absolute error.

[e]bpm: beats per minute.

[f]Number of data points.

In the overall sample (N=630), the correlation coefficient between CPX-measured HR and Fitbit Inspire 3–derived HR was 0.90 (95% CI 0.88 - 0.91; Figure 1). The MAE was 5.20 (SD 8.75) beats per minute (bpm), and the MAPE was 5.40% (SD 8.33%). The total error was 94 (N=630, 15%), with overestimation and underestimation of 37 (6%) and 57 (9%), respectively. The Bland-Altman plot (Figure 2) displays CPX-measured HR on the x-axis and the difference between CPX and Fitbit HR measurements on the y-axis. The time-series trend of HR error, with time on the x-axis, is shown in Figure 3. The average difference in HR was −1.25 bpm, with upper and lower limits of agreement of 18.56 bpm and −21.05 bpm, respectively.

**Figure 1.** Scatter plot comparing heart rate (HR) measured by the Fitbit Inspire 3 with HR measured by cardiopulmonary exercise testing (CPX), plotted on the x-axis. Each point represents an individual measurement and is color-coded by exercise phase (rest, below anaerobic threshold [AT], above AT, and recovery). The dashed gray line indicates the line of identity (y=x). bpm: beats per minute.

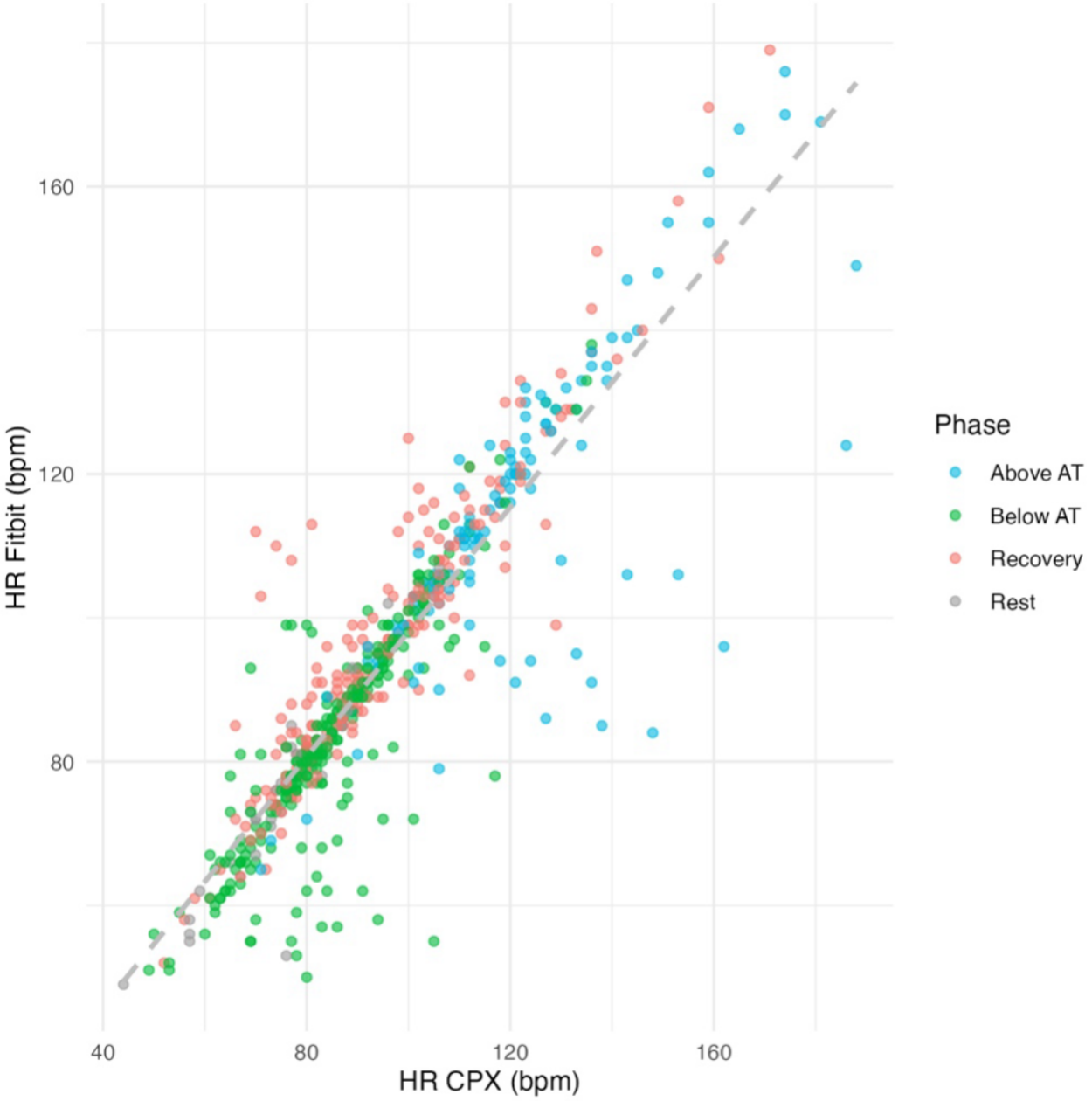


**Figure 2.** Bland-Altman plot comparing heart rate (HR) measurements from the Fitbit Inspire 3 and cardiopulmonary exercise testing (CPX). The x-axis represents HR measured by CPX, while the y-axis displays the difference between Fitbit and CPX HR values (Fitbit HR – CPX HR). The solid horizontal line indicates the mean bias (−1.25 beats per minute [bpm]), and the dashed lines represent the 95% limits of agreement (+18.56 bpm and −21.05 bpm). Density curves along the top and right margins show the distributions of CPX HR values and HR differences, respectively. Each point represents a single measurement and is color-coded by exercise phase (rest, below anaerobic threshold [AT], above AT, and recovery).

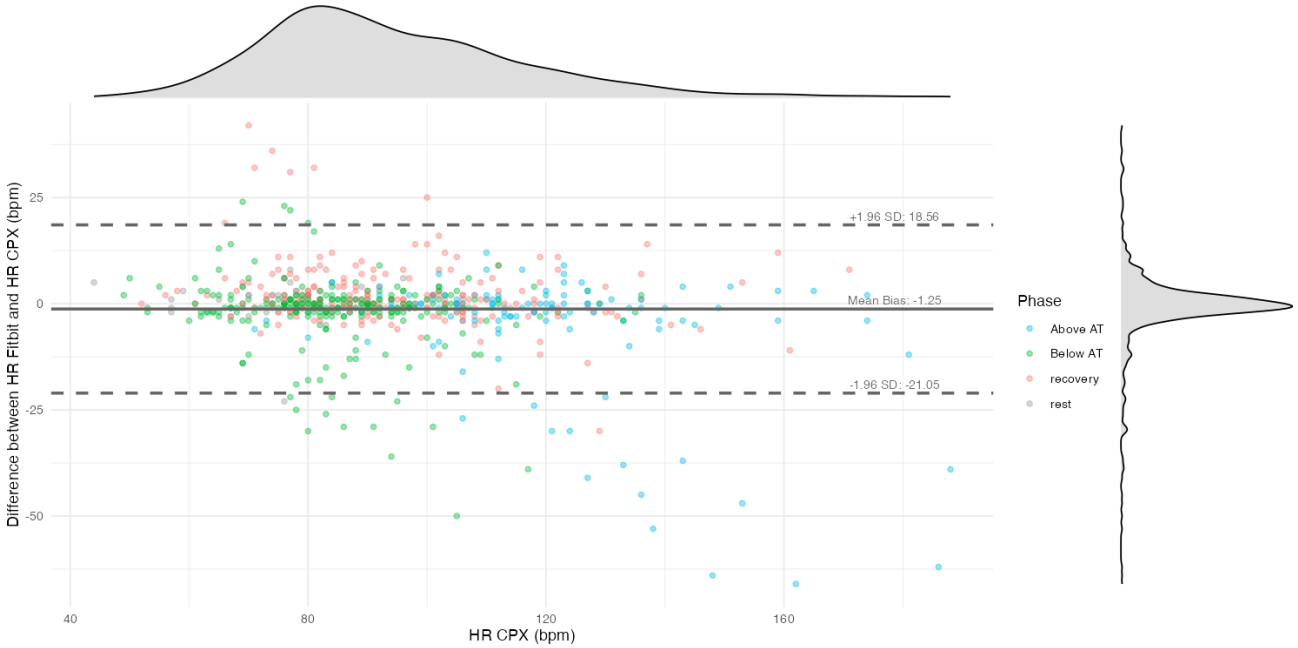
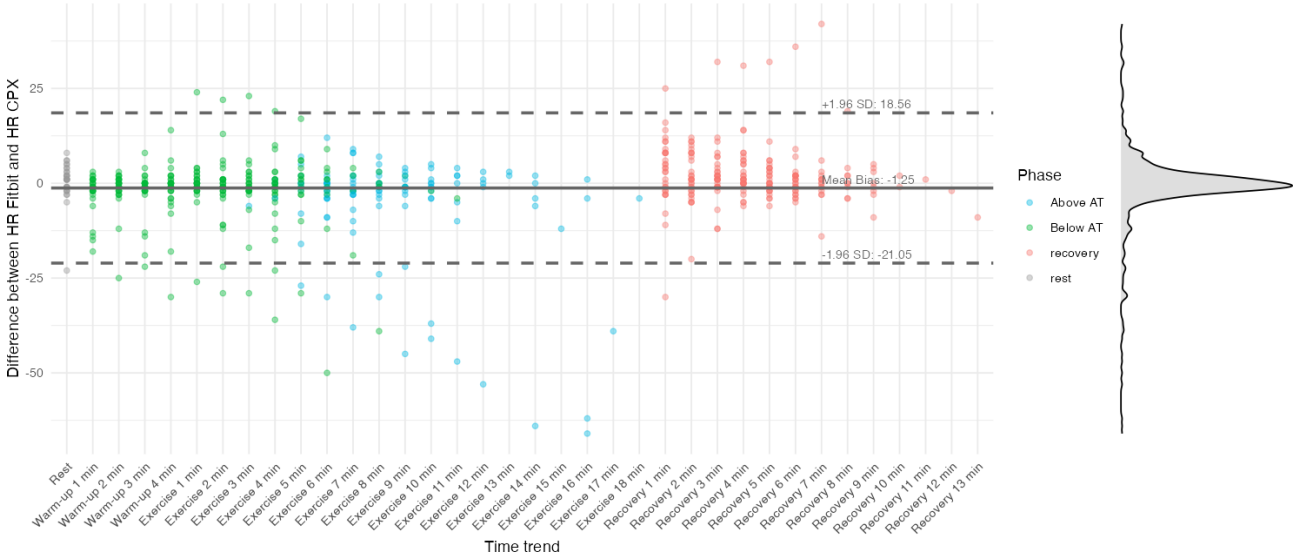



**Figure 3.** Time-series Bland-Altman plot comparing heart rate (HR) measurements from the Fitbit Inspire 3 and cardiopulmonary exercise testing (CPX). The x-axis represents time points across the CPX protocol, while the y-axis shows the difference in HR values (Fitbit HR – CPX HR). The solid horizontal line indicates the mean bias (−1.25 beats per minute [bpm]), and the dashed lines represent the upper (+18.56 bpm) and lower (−21.05 bpm) 95% limits of agreement. The density curve on the right illustrates the distribution of HR differences. Each point represents a single measurement, color-coded by exercise phase (rest, below anaerobic threshold [AT], above AT, and recovery).




By condition, the rest phase (n=30) showed a correlation coefficient of 0.85 (0.82 - 0.88), MAE of 4.41 (SD 7.28) bpm, and MAPE of 5.23% (SD 8.29%). During the below AT condition (n=270), the correlation coefficient was 0.92 (0.90 - 0.94), MAE was 4.85 (SD 6.46) bpm, and MAPE was 5.32% (SD 7.96%).

In contrast, the above AT condition (n=119) demonstrated a lower correlation coefficient of 0.76 (0.67 - 0.82), with higher MAE and MAPE values of 8.03 (SD 14.08) bpm and 6.14% (SD 9.62%), respectively. In the recovery phase (n=211), the correlation coefficient was 0.92 (0.90 - 0.94), with an MAE of 4.85 (SD 6.46) bpm and a MAPE of 5.32% (SD 7.96%). Regarding estimation errors, the above AT phase exhibited a higher underestimation (19/119 points, 15.9%) than overestimation (1/119 points; 0.8%), whereas in the recovery phase, overestimation (22/211 points, 10.4%) exceeded underestimation (5/211 points, 2.3%).

By HF stage, stage B (n=275) demonstrated a correlation coefficient of 0.92 (0.90 - 0.94) and an MAE of 4.51 (SD 7.66) bpm. In comparison, stage C (n=355) showed a slightly lower correlation coefficient of 0.88 (0.85 - 0.90) and a higher MAE of 5.73 (SD 9.49) bpm. The over- and underestimations were

17 (N=275, 6.1%) and 14 (5%) for stage B and 18 (N=355, 5%) and 40 (11.2%) for stage C, respectively.

## Discussion

### Principal Findings

In this study, we analyzed the accuracy of HR measurements obtained from the Fitbit Inspire 3 relative to ECG-based HR during CPX in patients with CVD, including those with HF. The Fitbit Inspire 3 device yielded relatively accurate HR estimates at intensities below the AT. However, estimation errors increased above the AT: the device underestimated HR during high-intensity exercise and overestimated it during recovery. In addition, overall accuracy was lower in patients with HF.

### Comparison With Previous Work

The overall correlation coefficient was 0.90 (Table 2), whereas previous studies involving healthy individuals using Fitbit devices reported correlation coefficients ranging from 0.84 to 0.93 [20,21]. MAPE was 4.67% (SD 5.53%) at rest and 6.14% (SD 9.62%) above the AT. A study using the Fitbit Charge 4 across various activities, such as stair climbing and squats, found MAPE values ranging from 6.36% to 11.98%, depending on the activity [22]. Another cycling-based study using the Fitbit Charge 3 reported a MAPE of 6.1% [23]. Despite using the comparatively less-expensive Fitbit Inspire 3 in this study, we observed similar outcomes, suggesting that, if the exercise modality is equivalent, similar accuracy may be achieved.

As the workload increased during CPX, MAPE rose from 4.67% (SD 5.53%) at rest to 5.23% (SD 8.29%) below the AT and 6.14% (SD 9.62%) above the AT (Table 2). Correspondingly, measurement error increased with HR (Figure 2). A previous study that varied walking speeds on a treadmill reported MAPEs of 9.99% at 3.0 km/h and 10.06% at 6.4 km/h [20]. Another study that varied cycling workloads, similar to this study, found a MAPE of −7.0% at a light load (50 W) and −15% at workloads of 60% - 85% of HR reserve [24], indicating larger measurement errors at higher exercise intensities. The decreased accuracy at elevated workloads may be attributable to several factors: the low test-retest reliability of wrist-based PPG sensors (intraclass correlation coefficients <0.5) [25], increased forearm muscle contraction to stabilize the handlebars (which can reduce blood flow and introduce signal artifacts) [26,27], reduced contact between the PPG sensor and the skin [28], measurement latency [29], and motion artifacts themselves [30]. In our study, underestimation became more pronounced at higher exercise intensities, whereas overestimation predominated during the recovery phase (Figure 3). The acute phase of HR recovery is influenced by parasympathetic reactivation, and the late phase is associated with sympathetic withdrawal and reduced catecholamine levels, generally leading to a 12 to 30 bpm decrease within 1 minute [31]. The fact that 23 (77%) participants were taking β blockers suggests that pharmacological effects, autonomic dysfunction, and the Fitbit device's delayed response to sudden changes in HR [24] may have collectively contributed to these discrepancies.

Notably, patients with stage C HF showed a higher rate of underestimation than those with stage B (11% vs 5%), likely owing to the larger measurement error at high workloads. No validation study has specifically focused on wearable devices in patients with HF. We speculate that because patients with HF have a limited ability to increase cardiac output beyond the AT [32] and experience reduced peripheral perfusion due to heightened sympathetic tone, PPG-based HR may be underestimated. Further research is warranted to clarify the accuracy of PPG-based HR measurements above the AT in this population.

### Limitations

This study has some limitations worth noting. First, it was a pilot study with a relatively small sample size. In addition, HR values were averaged over 1-minute intervals, leading to fewer data points. Consequently, rapid fluctuations in HR may not have been fully captured and could have contributed to measurement error. Second, we excluded patients with persistent AF; however, patients with both HF and AF may experience further reductions in HR measurement accuracy [33]. Third, we used cycling as the exercise modality, which restricted upper-limb movement. Home-based exercise therapies commonly involve walking, which requires arm movement. Therefore, we were unable to assess potential artifacts caused by arm motion. Further studies are needed to confirm the accuracy of these devices under more typical home-based exercise conditions. Fourth, the patients with HF in our cohort were relatively young and clinically stable. The results may differ in patients with more advanced HF, and caution should be exercised when generalizing these findings.

### Conclusions

This study demonstrated that the accuracy of HR estimation by the Fitbit Inspire 3 varied depending on exercise intensity and patient characteristics. These findings suggest that when using the Fitbit Inspire 3 to support interventions such as home-based exercise therapy in patients with CVD, including those with HF, careful consideration should be given to the patient's HF stage and exercise intensity, as well as to the device's potential limitations in different usage scenarios.

### Acknowledgments

The authors express their sincere gratitude to physical therapist Yuka Nakashima and the physical and occupational therapists in the Rehabilitation Department of Itami City Hospital for their invaluable support in facilitating this study. This work was supported by JSPS KAKENHI (grant JP24K20549).

## Data Availability

The data generated and analyzed during this research are available from the corresponding author upon reasonable request.

## Authors' Contributions

Conceptualization: KK
Data curation: KK (lead), YH (supporting)
Formal analysis: KK
Funding acquisition: KK
Investigation: YH, RF, TH
Methodology: KK
Project administration: KK (lead), YH (supporting)
Resources: YH, HM
Supervision: HS
Validation: KK
Visualization: KK
Writing–original draft: KK
Writing–review & editing: HS (lead), YH (supporting), RF (supporting)

## Conflicts of Interest

None declared.

## References

1. Makita S, Yasu T, Akashi YJ, et al. JCS/JACR 2021 Guideline on rehabilitation in patients with cardiovascular disease. Circ J 2022 Dec 23;87(1):155-235. [doi: 10.1253/circj.CJ-22-0234] [Medline: 36503954]
2. Fujimoto W, Toh R, Takegami M, et al. Estimating incidence of acute heart failure syndromes in Japan - an analysis from the KUNIUMI Registry. Circ J 2021 Sep 24;85(10):1860-1868. [doi: 10.1253/circj.CJ-20-1154] [Medline: 33678754]
3. Kamiya K, Yamamoto T, Tsuchihashi-Makaya M, et al. Nationwide survey of multidisciplinary care and cardiac rehabilitation for patients with heart failure in Japan - an analysis of the AMED-CHF Study. Circ J 2019 Jun 25;83(7):1546-1552. [doi: 10.1253/circj.CJ-19-0241] [Medline: 31189753]
4. Sugiharto F, Nuraeni A, Trisyani Y, Melati Putri A, Aghnia Armansyah N. Barriers to participation in cardiac rehabilitation among patients with coronary heart disease after reperfusion therapy: a scoping review. Vasc Health Risk Manag 2023;19:557-570. [doi: 10.2147/VHRM.S425505] [Medline: 37671387]
5. Molloy CD, Long L, Mordi IR, et al. Exercise-based cardiac rehabilitation for adults with heart failure - 2023 Cochrane systematic review and meta-analysis. Eur J Heart Fail 2023 Dec;25(12):2263-2273. [doi: 10.1002/ejhf.3046] [Medline: 37850321]
6. Sartor F, Vernillo G, de Morree HM, et al. Estimation of maximal oxygen uptake via submaximal exercise testing in sports, clinical, and home settings. Sports Med 2013 Sep;43(9):865-873. [doi: 10.1007/s40279-013-0068-3] [Medline: 23821468]
7. Kim C, Kim SH, Suh MR. Accuracy and validity of commercial smart bands for heart rate measurements during cardiopulmonary exercise test. Ann Rehabil Med 2022 Aug;46(4):209-218. [doi: 10.5535/arm.22050] [Medline: 36071003]
8. Friend SH, Ginsburg GS, Picard RW. Wearable digital health technology. N Engl J Med 2023 Nov 30;389(22):2100-2101. [doi: 10.1056/NEJMe2303219] [Medline: 38048193]
9. Fabien S, Waechter S, Kayode GA, Müller B, Roth M, Koehler F. Smartwatches in the assessment of heart failure patients in epidemiology and pathophysiology studies: a scoping review. ESC Heart Fail 2025 Jun;12(3):1727-1738. [doi: 10.1002/ehf2.15226] [Medline: 39905203]
10. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 8;8(9):e18694. [doi: 10.2196/18694] [Medline: 32897239]
11. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. J Sports Sci Med 2016 Aug 5;15(3):540-547. [Medline: 27803634]
12. Kim C, Song JH, Kim SH. Validation of wearable digital devices for heart rate measurement during exercise test in patients with coronary artery disease. Ann Rehabil Med 2023 Aug;47(4):261-271. [doi: 10.5535/arm.23019] [Medline: 37536665]
13. Helmer P, Hottenrott S, Rodemers P, et al. Accuracy and systematic biases of heart rate measurements by consumer-grade fitness trackers in postoperative patients: prospective clinical trial. J Med Internet Res 2022 Dec 30;24(12):e42359. [doi: 10.2196/42359] [Medline: 36583938]
14. Pierick AR, Burke KJ, Prusi M, et al. Validity of wrist-worn activity tracker heart rate detection in Fontan patients during exercise. Med Sci Sports Exerc 2025 Feb 1;57(2):280-284. [doi: 10.1249/MSS.0000000000003567] [Medline: 39294928]

15. Gorny AW, Liew SJ, Tan CS, Müller-Riemenschneider F. Fitbit Charge HR wireless heart rate monitor: validation study conducted under free-living conditions. JMIR Mhealth Uhealth 2017 Oct 20;5(10):e157. [doi: 10.2196/mhealth.8233] [Medline: 29055881]
16. Heidenreich PA, Bozkurt B, Aguilar D, et al. 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation 2022 May 3;145(18):e895-e1032. [doi: 10.1161/CIR.0000000000001063] [Medline: 35363499]
17. Bozkurt B, Coats AJS, Tsutsui H, et al. Universal definition and classification of heart failure: a report of the Heart Failure Society of America, Heart Failure Association of the European Society of Cardiology, Japanese Heart Failure Society and Writing Committee of the Universal Definition of Heart Failure: endorsed by the Canadian Heart Failure Society, Heart Failure Association of India, Cardiac Society of Australia and New Zealand, and Chinese Heart Failure Association, Heart Failure Association. Eur J Heart Fail 2021 Mar;23(3):352-380. [doi: 10.1002/ejhf.2115] [Medline: 33605000]
18. Boudreaux BD, Hebert EP, Hollander DB, et al. Validity of wearable activity monitors during cycling and resistance exercise. Med Sci Sports Exerc 2018 Mar;50(3):624-633. [doi: 10.1249/MSS.0000000000001471] [Medline: 29189666]
19. Nelson BW, Allen NB. Accuracy of consumer wearable heart rate measurement during an ecologically valid 24-hour period: intraindividual validation study. JMIR Mhealth Uhealth 2019 Mar 11;7(3):e10828. [doi: 10.2196/10828] [Medline: 30855232]
20. Stahl SE, An HS, Dinkel DM, Noble JM, Lee JM. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? BMJ Open Sport Exerc Med 2016;2(1):e000106. [doi: 10.1136/bmjsem-2015-000106] [Medline: 27900173]
21. Wang R, Blackburn G, Desai M, et al. Accuracy of wrist-worn heart rate monitors. JAMA Cardiol 2017 Jan 1;2(1):104-106. [doi: 10.1001/jamacardio.2016.3340] [Medline: 27732703]
22. Nissen M, Slim S, Jäger K, et al. Heart rate measurement accuracy of Fitbit Charge 4 and Samsung Galaxy Watch Active 2: device evaluation study. JMIR Form Res 2022 Mar 1;6(3):e33635. [doi: 10.2196/33635] [Medline: 35230250]
23. Haveman ME, van Rossum MC, Vaseur RME, et al. Continuous monitoring of vital signs with wearable sensors during daily life activities: validation study. JMIR Form Res 2022 Jan 7;6(1):e30863. [doi: 10.2196/30863] [Medline: 34994703]
24. Muggeridge DJ, Hickson K, Davies AV, et al. Measurement of heart rate using the Polar OH1 and Fitbit Charge 3 wearable devices in healthy adults during light, moderate, vigorous, and sprint-based exercise: validation study. JMIR Mhealth Uhealth 2021 Mar 25;9(3):e25313. [doi: 10.2196/25313] [Medline: 33764310]
25. Mol A, Meskers CGM, Niehof SP, Maier AB, van Wezel RJA. Pulse transit time as a proxy for vasoconstriction in younger and older adults. Exp Gerontol 2020 Jul 1;135(110938):110938. [doi: 10.1016/j.exger.2020.110938] [Medline: 32247853]
26. Johnson PC. Overview of the microcirculation. In: Tuma RF, Duran WN, Ley K, editors. Microcirculation: Academic Press; 2008:xi-xxiv. [doi: 10.1016/b978-0-12-374530-9.00022-x]
27. Johnson JM. Physical training and the control of skin blood flow. Med Sci Sports Exerc 1998 Mar;30(3):382-386. [doi: 10.1097/00005768-199803000-00007] [Medline: 9526883]
28. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 2007 Mar;28(3):R1-R39. [doi: 10.1088/0967-3334/28/3/R01] [Medline: 17322588]
29. Støve MP, Holm RS, Kjaersgaard AS, Duncker K, Jensen MR, Larsen BT. Measurement latency significantly contributes to reduced heart rate measurement accuracy in wearable devices. J Med Eng Technol 2020 Apr;44(3):125-132. [doi: 10.1080/03091902.2020.1753836] [Medline: 32404012]
30. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020;3(1):18. [doi: 10.1038/s41746-020-0226-6] [Medline: 32047863]
31. Petek BJ, Al-Alusi MA, Moulson N, et al. Consumer wearable health and fitness technology in cardiovascular medicine: JACC state-of-the-art review. J Am Coll Cardiol 2023 Jul 18;82(3):245-264. [doi: 10.1016/j.jacc.2023.04.054] [Medline: 37438010]
32. Nappi A, Cuocolo A, Imbriaco M, et al. Ambulatory monitoring of left ventricular function: walk and bicycle exercise in congestive heart failure. J Nucl Med 1997 Jun;38(6):948-953. [Medline: 9189148]
33. Al-Kaisey AM, Koshy AN, Ha FJ, et al. Accuracy of wrist-worn heart rate monitors for rate control assessment in atrial fibrillation. Int J Cardiol 2020 Feb 1;300:161-164. [doi: 10.1016/j.ijcard.2019.11.120] [Medline: 31787389]


## Abbreviations

**ACC:** American College of Cardiology
**AF:** atrial fibrillation
**AHA:** American Heart Association
**AT:** anaerobic threshold
**bpm:** beats per minute
**BW:** body weight
**CAD:** coronary artery disease
**CPX:** cardiopulmonary exercise testing
**CR:** cardiac rehabilitation

**CVD:** cardiovascular disease
**ECG:** electrocardiography
**HF:** heart failure
**HR:** heart rate
**LVEF:** left ventricular ejection fraction
**MAE:** mean absolute error
**MAPE:** mean absolute percentage error
**PPG:** photoplethysmography
**$VCO_2$:** carbon dioxide output
**VE:** minute ventilation
**$VO_2$:** oxygen uptake

*Edited by G Krstačić; submitted 22.05.25; peer-reviewed by M Støve, V Damasceno; revised version received 04.08.25; accepted 03.09.25; published 06.10.25.*

*Please cite as:*
*Kitagaki K, Hongo Y, Futai R, Hasegawa T, Morikawa H, Shimoyama H*
*Validity of Heart Rate Measurement Using Wearable Devices During Cardiopulmonary Exercise Testing in Patients With Cardiovascular Disease: Prospective Pilot Validation Study*
*JMIR Cardio 2025;9:e77911*
*URL: https://cardio.jmir.org/2025/1/e77911*
*doi:10.2196/77911*

# Wrist-Worn and Arm-Worn Wearables for Monitoring Heart Rate During Sedentary and Light-to-Vigorous Physical Activities: Device Validation Study

Theresa Schweizer, MSc; Rahel Gilgen-Ammann, PhD

Department of Monitoring and Evaluation, Swiss Federal Institute of Sport Magglingen SFISM, Hauptstrasse 247, Magglingen, Switzerland

**Corresponding Author:**
Theresa Schweizer, MSc
Department of Monitoring and Evaluation, Swiss Federal Institute of Sport Magglingen SFISM, Hauptstrasse 247, Magglingen, Switzerland

## Abstract

**Background:** Heart rate (HR) is a vital physiological parameter, serving as an indicator of homeostasis and a key metric for monitoring cardiovascular health and physiological responses. Wearable devices using photoplethysmography (PPG) technology provide noninvasive HR monitoring in real-life settings, but their performance may vary due to factors such as wearing position, blood flow, motion, and device updates. Therefore, ongoing validation of their accuracy and reliability across different activities is essential.

**Objectives:** This study aimed to assess the accuracy and reliability of the HR measurement from the PPG-based Polar Verity Sense and the Polar Vantage V2 devices across a range of physical activities and intensities as well as wearing positions (ie, upper arm, forearm, and both wrists).

**Methods:** Sixteen healthy participants were recruited to participate in this study protocol, which involved 9 activities of varying intensities, ranging from lying down to high-intensity interval training, each repeated twice. The HR measurements from the Verity Sense and Vantage V2 were compared with the criterion measure Polar H10 electrocardiogram (ECG) chest strap. The data were processed to eliminate artifacts and outliers. Accuracy and reliability were assessed using multiple statistical methods, including systematic bias (mean of differences), mean absolute error (MAE) and mean absolute percentage error (MAPE), Pearson product moment correlation coefficient ($r$), Lin concordance correlation coefficient (CCC), and within-subject coefficient of variation (WSCV).

**Results:** All 16 participants (female=7; male=9; mean 27.4, SD 5.8 years) completed the study. The Verity Sense, worn on the upper arm, demonstrated excellent accuracy across most activities, with a systematic bias of −0.05 bpm, MAE of 1.43 bpm, MAPE of 1.35%, $r$=1.00, and CCC=1.00. It also demonstrated high reliability across all activities with a WSCV of 2.57% and no significant differences between the 2 sessions. The wrist-worn Vantage V2 demonstrated moderate accuracy with a slight overestimation compared with the ECG and considerable variation in accuracy depending on the activity. For the nondominant wrist, it demonstrated a systematic bias of 2.56 bpm, MAE of 6.41 bpm, MAPE 6.82%, $r$=0.93, and CCC=0.92. Reliability varied considerably, ranging from a WSCV of 3.64% during postexercise sitting to 23.03% during lying down.

**Conclusions:** The Verity Sense was found to be highly accurate and reliable, outperforming many other wearable HR devices and establishing itself as a strong alternative to ECG-based chest straps, especially when worn on the upper arm. The Vantage V2 was found to have moderate accuracy, with performance highly dependent on activity type and intensity. While it exhibited greater variability and limitations at lower HR, it performed better at higher intensities and outperformed several wrist-worn devices from previous research, particularly during vigorous activities. These findings highlight the importance of device selection and wearing position to ensure the highest possible accuracy in the intended context.

***(JMIR Cardio 2025;9:e67110)*** doi:10.2196/67110

**KEYWORDS**

validity; reliability; accuracy; wearable devices; wearing position; photoplethysmography; heart rate

## Introduction

Heart rate (HR) is one of the most commonly measured physiological parameters in wearables, valued for its ease of measurement and its role as a key marker of homeostasis, cardiovascular health, and physiological responses. HR can provide early warnings for certain pathological conditions; for example, resting HR is an independent predictor of cardiovascular disease, stroke, and sudden death [1,2]. In addition, HR is frequently used for assessing physical effort,

workload intensity, and supporting performance monitoring. It is also often integrated into algorithms to estimate other physiological metrics, such as core body temperature and energy expenditure [3-5]. HR is therefore a valuable and valid parameter when aiming for health monitoring and workload management.

The current criterion measure for assessing HR outside the laboratory is the chest strap, which uses electrocardiogram (ECG) technology, due to its strong agreement and minimal bias when compared with the ECG-Holter device in healthy adults and patients [6-10]. A prior validation study demonstrated that the Polar H10 (H10; Polar Electro Oy) exhibited even higher accuracy during higher-intensity activities with increased motion than the ECG-Holter [11]. However, the continuous use of chest straps every day in the field can lead to discomfort, incompatibility with equipment, or displacement issues [12]. Consequently, there is growing interest in wrist-, upper arm-, or forearm-wearable devices, which use photoplethysmography (PPG) [13]. PPG is a noninvasive measurement technique that detects blood volume changes in the microvascular bed of tissue by illuminating the skin and measuring the reflected light [14].

The affordability and capability of these wearable devices to continuously monitor physiological parameters over extended periods, combined with rapid advancements in multimodal sensing technologies and extensive marketing by manufacturers, have led to their widespread use. However, the quality of the data is crucial when monitoring health parameters in real life. Many users—and even scientists—may rely on these devices to measure outcomes such as resting HR, training zones, fatigue, or health issues without verifying the accuracy and reliability of the measured physiological parameters. Notably, one critical review showed that more than half of the technologies reviewed had not been validated through independent research, with only 5% having been formally validated [13]. As wearable technologies continue to evolve with each update or new version including new sensor modalities, it is important to conduct ongoing assessements of their accuracy and reliability, as these factors can impact measurement performance [1,15-18].

Furthermore, validation studies often focus on only 1 or a few standardized exercises (eg, resting, cycling, or treadmill running) that involve minimal movement artifacts in the arms or wrists and are conducted in controlled laboratory settings [19-21]. In fact, HR measurement accuracy has shown to be influenced by differences in blood flow, motion artifacts, and the interaction between the sensor and skin on the different wearing position [22-25]. For example, proximal wearing position such as the upper arm may provide more stable readings during high-motion activities than distal placements such as the forearm or the wrist, where movement artifacts are more pronounced and blood flow is lower. For HR monitoring to be applicable to general activity tracking, data should be validated across a variety of exercise modalities at different intensities (resting, submaximal, and high) and body positions (lying, sitting, and standing), as well as during free movement [15].

Although the H10 is recognized as a criterion measure based on the INTERLIVE Network's expert statement [26], the Polar Verity Sense (Polar Electro Oy) offers a possible alternative. When worn on the upper arm, the Verity Sense sits well on the skin, may be less intrusive than a chest strap, and provides advantages over a wrist-worn device due to its proximal wearing position (eg, increased blood flow). The Verity Sense has been evaluated in prior studies, though the activities were in some of the studies very short, laboratory-based, in paced conditions, or very specific (eg, walking, jogging, swimming, Pickleball Game Play, or biking) [27-31]. Similarly, the Vantage V2 has been validated in prior studies, but the studies had either an older criterion measure or was validated in specific activities in laboratory conditions (eg, paced running and swimming) [31-33]. To the authors' knowledge, no study has evaluated the different wearing locations and tested it in various types of exercises and intensities in a more naturalistic environment.

Therefore, this study aims to validate the Polar Verity Sense and Vantage V2 in terms of HR across diverse activities, intensities, and wearing positions in conditions that closely resemble free-living environments over a sufficient amount of time to get robust results. The study incorporates a variety of activities, including different resting (eg, lying and sitting), common exercises (eg, running and cycling), body weight exercises, and dynamic movements such as parkour, which introduce significant challenges such as variations in blood flow and involve high levels of motion. To ensure robust findings, the protocol will be repeated twice to assess the reproducibility of HR measurements.

## Methods

### Participants

Sixteen healthy participants were recruited for this study. Recruitment was conducted via email announcements and in-person assessments of students and staff at the Swiss Federal Institute of Sport Magglingen. The study aimed to include individuals with diverse fitness levels and training habits, ensuring representation of both those who met and those who did not meet the World Health Organization's recommendation of 150 - 300 minutes of moderate-intensity aerobic physical activity per week [34]. Participants had to be between 18 and 40 years of age with a BMI between 18.5 and 30 kg/m². Interested participants received detailed study information and provided written informed consent before participation. Prior to inclusion, they were screened using the Physical Activity Readiness Questionnaire to ensure that they met the eligibility criteria. Only those who answered "no" to all Physical Activity Readiness Questionnaire questions, did not take any medication affecting HR, had no known ECG abnormalities, and had no tattoos on the sensor placement areas (upper arms, forearms, and wrists) were included in the study. In addition, skin type was assessed using the Fitzpatrick Scale [35], and the amount of body hair on the wrists and arms was recorded.

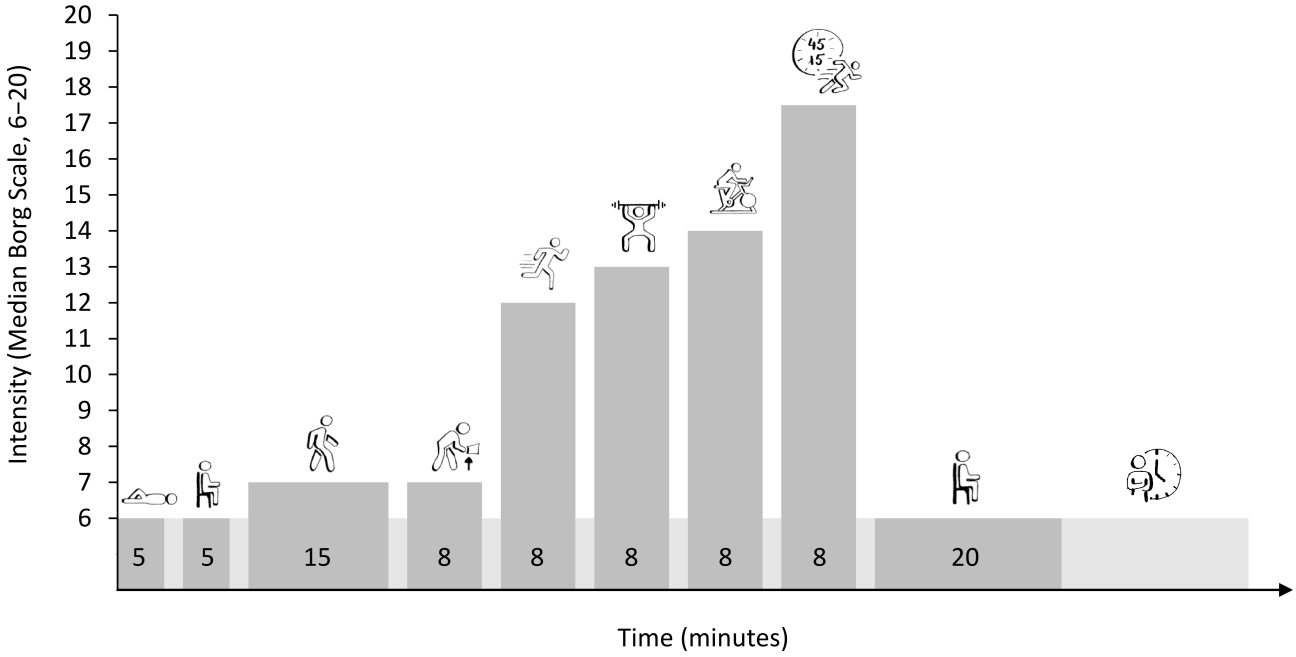
### Experimental Procedure

The participants were tested individually on different days and at different times of the day. The measurements were conducted in a gymnasium with prepared areas to perform the different activities and with consistent environmental conditions, with a mean (SD) ambient temperature of 19.5 °C (SD 0.9 °C) and humidity of 49.8% (SD 3.9%). After recording each participant's

weight, height, skin color, and body hair (while they were dressed in underwear), all devices were placed in the specific wearing positions on the body as recommended by the manufacturers. The H10 chest strap was moistened prior to use. All devices were activated at least 5 minutes before the protocol began to allow the sensors to calibrate to the HR.

The study protocol consisted of 9 different activities in order of increasing intensity (Figure 1): lying down (5 minutes), sitting (5 minutes), walking (15 minutes), picking up objects (8 minutes), jogging (8 minutes), weight training (8 minutes consisting of squats, biceps curls, lunges, and abdominal crunches), cycling on an ergometer (8 minutes), high-intensity interval training (HIIT; 8 minutes of a continuous parkour containing sprinting, dragging, carrying, lifting, and hammering, with 45 seconds of effort and 15 seconds of rest), and postexercise sitting (20 minutes). A 2-minute rest was taken between activities, and the entire protocol was repeated twice, with a 20-minute break between sessions in which the participants sat down, rested, and could drink or eat something, if needed. The procedures and instructions were standardized and identical for all participants, but they were kept very short to enhance the naturalistic study design. The participants rated their exertion using the Borg Rating of Perceived Exertion scale (6 - 20) after each activity to quantify intensity levels, ranging from minimal to near-maximal exertion [36,37].

**Figure 1.** Study protocol with 9 activities with 2-minute breaks in between. This protocol was repeated twice with a 20-minute break between sessions. Lower-intensity activities, such as lying down, sitting, and postexercise sitting, showed a median (IQR) rating of perceived exertion (RPE) of 6.0 (1.0), indicating minimal exertion. Low-intensity activities, including walking and picking up objects, had RPE values of 7.0 (1.25) and 7.0 (2.0), respectively, while jogging and weight training had RPE values of 12.0 (2.25) and 13.0 (2.0). Higher-intensity activities, such as cycling and high-intensity interval training, had median RPEs of 14.0 (2.25) and 17.5 (2.0), respectively, the latter reflecting near-maximum exertion. Across all activities, the median RPE was 10.0 (7.0).




## Devices and Instruments

### *Wearable Devices*

The Polar H10 (H10) measures HR using 1-lead ECG technology with a sampling frequency of 1000 Hz. According to the INTERLIVE Network's expert statement, ECG chest straps that have been independently validated and demonstrate excellent agreement with respect to beats per minute (ie, >95%) are considered appropriate criterion measures for evaluating wearable technologies measuring HR [26]. The H10 is included in their list of validated devices, with a prior study showing an excellent agreement ($r$=0.997) and 97.1% of the measured RR intervals (ie, time between successive R-wave peaks in the QRS complex—a waveform in an ECG representing ventricular depolarization and contraction, which corresponds to one full cardiac cycle) differing by less than 2% during various activities and intensities [11].

In this study, 2 wearable devices were evaluated. Both were placed on different wearing positions. The Verity Sense (Polar Electro Oy) measures HR on the upper arm and forearm using optical PPG technology with a sampling frequency of 1 Hz (firmware version: 2.0.3). The Vantage V2 (Polar Electro Oy) measures HR on the wrist using optical PPG technology with a sampling frequency of 1 Hz (firmware version: 4.1.0). Figure 2 shows the devices included in the study as well as their positions on the body. The Verity Sense devices were placed on the forearm and upper arm of opposite sides, with the specific side (left or right) randomly assigned across participants. Two Vantage V2 watches were placed on the wrists of each participant to capture readings from both the dominant and nondominant sides. One more Vantage V2 was used as a data

logger for the H10 and placed in a small pocket on an elastic belt around the waist. The Vantage V2 were started in the activity mode “other indoor” as no Global Positioning System was needed and different activities were performed. The Verity Sense were started in “recording mode”. All data were downloaded from the web-based Polar Flow application (Polar Electro Oy).

**Figure 2.** Placement of the different wearable devices. The H10 chest belt was placed on the chest with a Vantage V2 as logger on the waist. A Vantage V2 was placed on each wrist. A Verity Sense was placed on the upper arm and forearm.

### *Other Instruments*

The body heights of the participants were measured using a stadiometer (model 214; Seca GmbH), and body weight was measured on a calibrated digital balance scale (model 877; Seca GmbH). The cycling ergometer Ergoselect 200 (Ergoline GmbH) was used for the cycling activity, and dumbbells weighing from 2.5 to 10 kg were used for the weight training. A weather station was used to measure ambient temperature and humidity.

### Data Processing and Cleaning

First, all rest periods between activities were removed from the data. Second, the HR data derived from the PPGs (Verity Sense and Vantage V2) were synchronized with the reference using time stamps from the exported file and cross-correlated to fix the inconsistent lags between the ECG- and PPG-derived HR signals [38,39]. Third, missing values (ie, blanks or zeros) and artifacts were quantified. Data were considered artifacts if they fell below 30 bpm (type I), if they exceeded 230 bpm (type II), or if consecutive values differed by 15 bpm (type III) [40,41]. All artifacts were then removed from the dataset. Fourth, all reference data from the H10 device were statistically and visually inspected for potential outliers or irregularities to prevent errors from being mistakenly attributed to the Verity Sense and Vantage V2 devices. For each participant, the activities were flagged if they contained more than 10 missing data points, more than 10 artifacts, or a Pearson correlation below 0.9 compared with the Verity Sense or Vantage V2. The flagged activities underwent further visual screening to identify whether the error originated from the H10. If the H10 data contained a substantial number of outliers or were considered irregular, the entire activity was excluded from the analysis. Finally, HR data were averaged in 10-second intervals for each activity.

### Statistical Analysis

Statistical analysis was performed in accordance with previous recommendations [15]. The data from the tested devices and the criterion measure were assessed for normality, and all data were found to be normally distributed.

Accuracy was assessed for overall data and for each activity using systematic bias (mean of differences) with 95% limits of agreement (LoA), accompanied by the results of a 2-tailed 1-sample *t* test performed on the differences between the 2 measurements (ie, difference from zero). Moreover, mean absolute error, mean absolute percentage error (MAPE), 5% accuracy (percentage of MAPE within a 5% range of the reference value), root-mean-squared error (RMSE), and ordinary least squares linear regression were used to evaluate accuracy. Although previous validation studies lack consensus and have defined varying accuracy thresholds, this study classified a device as having very high accuracy if MAPE was <3%, high accuracy if MAPE was <5%, and moderate accuracy if MAPE was <10%, based on criteria used in some validation studies [21,28,31,42,43]. Pearson product moment correlation coefficient (*r*) and Lin concordance correlation coefficient (CCC) were used to evaluate the agreement between the criterion measure and the wearable device [44-46]. The Pearson correlation coefficient was interpreted as follows: 0.45 - 0.69 (very poor), 0.70 - 0.84 (poor), 0.85 - 0.94 (good), 0.95 - 0.994 (very good), and >0.995 (excellent) [47]. The strength-of-agreement criteria for the CCC were interpreted using McBride's (2005) criteria: <0.90 (poor agreement), 0.90 - 0.95 (moderate agreement), 0.95 - 0.99 (substantial agreement), and >0.99 (almost perfect agreement) [44].

Reliability was assessed using the within-subject coefficient of variation (WSCV), calculated based on the differences between the tested devices and the reference data, where lower values indicate greater consistency. Based on a prior study, the threshold of <5% was used to indicate high reliability, while <10% was considered acceptable reliability [21]. In addition, reproducibility was assessed using the Wilcoxon signed rank test to compare the differences between the device and reference measurements between session 1 and session 2. All data processing, cleaning, and analysis was done with Python (version 3.12; Python Software Foundation).

### Ethical Considerations

This study involving human participants was reviewed and approved by the Swiss ethics committee (project ID: 2022 - 01456). The research design adhered to the ethical standards outlined in the Declaration of Helsinki. All data collected were deidentified to ensure participant confidentiality. No personal identifiers were included in the dataset, and access to raw data was restricted to authorized researchers only. Participants provided written informed consent, which included permission for their anonymized data to be used in publications and shared with other researchers for further research purposes, in strict adherence to data protection regulations. Participants received a gift card valued at 30 Swiss Francs (CHF), approximately US $29 based on the exchange rate at the time of the study, as compensation for their time and participation. No identifiable images of participants are included in the manuscript or supplementary materials.

## Results

### Participants

Sixteen healthy participants (female=7; male=9; dominant right-handed=13) volunteered for this study. Their demographic characteristics reported as mean (SD) were age: 27.4 (5.8) years, height: 173.5 (9.2) cm, weight: 69.9 (9.4) kg, and BMI: 23.1 (2.0) kg/m$^2$. Ten participants met the recommendations of the World Health Organization of 150 - 300 minutes of moderate-intensity aerobic physical activity per week and 6 were below that threshold. Six participants were classified as type I, and 10 participants were classified as type II according to the Fitzpatrick Scale. In addition, none of the participants had exceptionally hairy skin at any of the device-wearing positions.

### Missing Values, Artifacts, and Outliers

No devices had missing values; however, artifacts and outliers were identified in the H10 and Verity Sense data. For the H10, 9 randomly occurring type III artifacts were found. In addition, visual screening led to the overall removal of 16,462 seconds (10%) of the raw data from 3 participants, including the entire

protocol's first session of 1 participant and the second session of 2 participants. These outliers were potentially due to suboptimal positioning or displacement of the H10 in these 3 participants. In the Verity Sense data, 85 seconds (0.06%) were classified as type I artifacts (upper arm: 36; forearm: 49) and 32 seconds (0.02%) as type III artifacts (upper arm: 3; forearm: 29). No specific activity, participant, or gender could be identified as having more artifacts than the others.

After averaging the cleaned data into 10-second intervals, the data from the 16 participants totaled 40.7 hours (mean 4.5, SD 2.1 hours per participant), resulting in 14,653 10-second data points analyzed across all activities. The sedentary or resting activities, including lying down, sitting, and postexercise sitting, contributed 867, 870, and 3346 data points, respectively, totaling 5083 (34.7%) data points. Low- to moderate-intensity activities, such as walking and picking up objects, provided 2610 and 1392 data points, respectively, amounting to 4002 (27.3%) data points. Higher-intensity activities, including jogging, weight training, cycling, and HIIT, each contributed 1392 data points, for a total of 5568 (38.0%) data points. This distribution ensured comprehensive coverage across all activity types and intensities.

## Accuracy and Reliability

### *Arm-Worn Verity Sense*

The overall mean bias was −0.05 bpm (LoA –5.84 to 5.74 bpm) on the upper arm and −0.91 bpm (LoA –14.64 to 12.83) on the forearm, indicating only minimal underestimation of the HR measurements. The 2-tailed 1-sample *t* test was conducted to determine whether the differences between the Verity Sense and the reference measurement significantly deviated from zero. The results indicated no significant difference on the upper arm for lying ($P$=.845), sitting ($P$=.093), jogging ($P$=.159), and postexercise sitting ($P$=.911). Likewise, on the forearm, no significant differences were found for lying ($P$=.981), walking ($P$=.227), and jogging ($P$=.306). No significant differences were found overall and for all other activities ($P$<.05). For the upper arm placement, MAPE remained low across all activities, with the lowest values observed during jogging (0.69%) and cycling (0.53%) and the highest during sitting (2.48%) and picking up objects (2.34%). On the forearm, MAPE was slightly higher overall, with the lowest values recorded during jogging (0.92%) and cycling (0.60%). The overall 5% accuracy was 95% for the upper arm and 89% for the forearm. The RMSE for the upper arm was generally low across activities, with an overall value of 2.95 bpm, except for weight training, which showed an RMSE of 6.49 bpm. RMSE values for the forearm were higher, with an overall mean of 7.07 bpm. Pearson correlation coefficients demonstrated very good to excellent positive linear correlations between the Verity Sense and the ECG criterion across all activities for the upper arm ($r$>0.94). For the forearm, the correlations similarly ranged from very good to excellent for all activities ($r$>0.95), except weight training ($r$>0.88), HIIT ($r$>0.85), and postexercise sitting ($r$>0.79). Regression analyses supported these findings, with strong correlations ($r^2$=0.99 for the upper arm and $r^2$=0.96 for the forearm) and regression slopes near 1.00, especially during lower-intensity activities, except for weight training. The CCC showed consistently almost perfect agreement, with an overall CCC of 1.00 (95% CI 0.99-1.00) for the upper arm, although lower values were observed during weight training. For the forearm, the CCC showed substantial agreement with an overall value of 0.98 (95% CI 0.97-0.98), with decreased agreement during HIIT and postexercise sitting.

The Verity Sense demonstrated high reliability across most activities, regardless of arm placement. The Wilcoxon signed rank test showed no significant differences between the device and reference measurements across sessions for the upper arm ($W$=2994.0, $P$=.213; session 1: $mean_{diff}$ –0.14 bpm, $SD_{diff}$ 0.87 bpm; session 2: $mean_{diff}$ –0.07 bpm, $SD_{diff}$ 1.70 bpm) and forearm ($W$=3081.0, $P$=.314; session 1: $mean_{diff}$ –0.61 bpm, $SD_{diff}$ 2.63 bpm; session 2: $mean_{diff}$ –1.06 bpm, $SD_{diff}$ 5.74 bpm) placements. In addition, the WSCV was consistently low, particularly for the upper arm (ranging from 0.98% for cycling to 4.98% for weight training), while the forearm exhibited slightly higher variability (1.14% for cycling to 9.80% for postexercise sitting).

Table S1 in Multimedia Appendix 1 shows the detailed accuracy and reliability results for the Verity Sense compared with the reference for each activity and for each wearing position.

### *Wrist-Worn Vantage V2*

The overall mean bias was 2.93 bpm (LoA –20.46 to 26.31) and 2.56 bpm (LoA –21.88 to 26.99) for the dominant and nondominant wrists, respectively, indicating a slight overestimation of HR with large LoAs. For the 2-tailed 1-sample *t* test, for both the dominant and nondominant wrists, no significant difference was found for sitting ($P$=.271; $P$=.818), whereas all other activities showed significant differences ($P$<.001).

For both wearing positions (dominant and nondominant), MAPE was lowest during jogging (3.84% and 3.55%), cycling (1.17% and 2.06%), and postexercise sitting (2.15% and 2.07%). However, MAPE exceeded 10% during activities characterized by lower HR, such as lying down, walking, and picking up objects. The 5% accuracy showed varying levels of agreement across all activities, with an overall result of 73.56% for the dominant wrist and 71.83% for the nondominant wrist. For both the dominant and nondominant wrists, RMSE was generally high, with overall values of 12.29 bpm and 12.73 bpm, respectively. However, accuracy improved during postexercise sitting, where RMSE was lower at 3.60 bpm and 3.78 bpm. Pearson correlation and regression analyses further highlighted these discrepancies. For both the dominant and nondominant wrists, correlation was good to very good during jogging ($r$=0.89 and $r$=0.91), weight training ($r$=0.90 and $r$=0.91), cycling on an ergometer ($r$=0.98 and $r$=0.94), and postexercise sitting ($r$=0.97 and $r$=0.97). However, accuracy was very poor to poor for all other tasks. A slight difference between wearing positions was observed during HIIT, where the dominant wrist showed poor correlation ($r$=0.81), while the nondominant wrist showed good correlation ($r$=0.85). In addition, linear regression slopes indicated overall low agreement, with values of 0.87 and 0.85 for the dominant and nondominant wrists, respectively. On the dominant wrist, CCC ranged from poor agreement (0.25 during picking up objects) to substantial agreement (0.97 during cycling). On the nondominant wrist, CCC values ranged from

poor agreement (0.24 during picking up objects) to substantial agreement (0.97 during postexercise sitting).

The Vantage V2 demonstrated moderate reliability across most activities for both wrist placements. The Wilcoxon signed rank test showed no significant differences between the device and reference measurements across sessions for the dominant wrist ($W$=3379.0, $P$=.844; session 1: $mean_{diff}$ 3.72 bpm, $SD_{diff}$ 10.96 bpm; session 2: $mean_{diff}$ 3.63 bpm, $SD_{diff}$ 10.32 bpm) and the nondominant wrist ($W$=2852.5, $P$=.103; session 1: $mean_{diff}$ 3.51 bpm, $SD_{diff}$ 12.37 bpm; session 2: $mean_{diff}$ 2.41 bpm, $SD_{diff}$ 8.73 bpm). Although no significant differences were found between sessions, the WSCV varied across activities. Lower variability was observed for postexercise sitting (3.49% on the dominant wrist; 3.64% on the nondominant wrist), while very high variability was found during lying down (26.44% on the dominant wrist; 23.04% on the nondominant wrist). Overall, variability remained high, with overall WSCV values of 10.41% for the dominant wrist and 10.87% for the nondominant wrist.

Table S1 in Multimedia Appendix 2 shows the detailed accuracy and reliability results for the Vantage V2, compared with the reference for each activity and for each wrist placement.

# Discussion

## Principal Findings and Comparison With Prior Work

### *Arm-Worn Polar Verity Sense*

This study evaluated the accuracy and reliability of the arm-worn Verity Sense across various activities and both placements, the forearm and the upper arm. The device had no missing values and only a trivial number of artifacts (0.08%). Overall, and especially on the upper arm, the Verity Sense demonstrated minimal bias (−0.05 bpm), very high accuracy (MAPE 1.35%), and very good to excellent agreement with ECG ($r$=1.00, CCC 1.00). Reliability was also high, with no significant differences between sessions and consistently low variability in comparison with the criterion measure (WSCV 2.57%).

The overall trend suggested the highest accuracy and reliability during activities with elevated mean HR and less arm movements, while slightly lower accuracy was noted during low-intensity tasks such as weight training and object picking. As PPG-based HR measurements are influenced by differences in blood flow and motion artifacts, these findings underline the possible loss of accuracy with increased motion as well as reduced lower blood flow (eg, lower HR, cold extremities, and blood flow restriction due to clothes or other devices) [22-25]. These results align with previous studies that reported reduced accuracy in similar low-intensity, high-motion activities [16,28,31]. Notably, even during these challenging tasks, the upper arm placement continued to deliver strong results.

To the authors' knowledge, regardless of the wearing position on the upper arm or the forearm, the excellent accuracy demonstrated by the Verity Sense in this study outperformed all of the following wearable devices tested in different activities and settings in previous studies: multiple Garmin wrist-worn devices (eg, Instinct, Venu, and Fenix 5 - 6) [20,27,28,32,33,48,49], various Polar wrist-worn devices and the OH1 (ie, the prior version of the Verity Sense) [21,27,28,30,32,48], the Apple Watch [20,49], the Motiv Ring, the arm-worn Scosche Rythm+, the Jabra Elite Sport and the Suunto Spartan Sport [20], FitBit Charge 2 and 4 [19,43,50], and the Samsung Galaxy Watch Active2 [43].

In addition, in this study, the Verity Sense outperformed its own previous results from studies conducted between 2022 and 2024, demonstrating better MAPE values while maintaining similar regression analysis and CCCs [27-31,48]. These results suggest that the Verity Sense is a highly accurate and reliable alternative to the ECG-based chest strap such as the Polar H10. Notably, given the number of missing values and artifacts observed in the H10 in this study, the Verity Sense may offer greater robustness across the investigated activities. However, this study does not provide conclusive evidence of interchangeability between these devices.

### *Wrist-Worn Polar Vantage V2*

This study evaluated the accuracy and reliability of the wrist-worn Vantage V2 across various activities and both wrist placements (dominant and nondominant). The device had no missing values or artifacts, suggesting a robust filtering method, as wrist-worn devices typically experience significant motion artifacts and low blood flow [22-25]. The Vantage V2 performed similarly on both wrists, showing a slight HR overestimation with large LoAs and overall moderate accuracy. However, accuracy varied considerably depending on the activity. High accuracy (MAPE<5%) was observed in all moderate- to vigorous-intensity activities (ie, jogging, weight training, cycling, and HIIT) as well as postexercise sitting, whereas activities with lower HR and increased motion artifacts exhibited poorer accuracy. Overall, although CCC demonstrated moderate agreement, Pearson correlation indicated good agreement and reached very good agreement during cycling on an ergometer and postexercise sitting, the 2 activities with low arm and wrist movement as well as increased blood flow. However, it is important to note that high correlations do not guarantee the absence of bias or error, nor do they confirm perfect validity [51]. Although no significant differences between sessions were found, overall reliability was below the acceptable threshold, with WSCVs exceeding 10%. Variability was particularly high during low-intensity activities (eg, lying down and picking up objects). In contrast, high to very high reliability was observed again during cycling on an ergometer and postexercise sitting. This again highlights the influence of motion artifacts combined with lower HR (ie, blood flow) on signal quality at the wrist position.

In previous studies, wrist-worn devices showed similar results: the bias tends to increase with the intensity of activity on a treadmill, while using a cycle ergometer, and during resistance training tasks [19,42,48,49,52,53]. Similarly, one study found that the magnitude of the errors depended on the activity type and that it can result in an absolute error that is 30% higher than at rest [38]. Wrist-worn devices are more susceptible to noise and distortion due to thinner skin, underlying bones and tendons, and reduced blood perfusion, all of which increase the likelihood of motion artifacts in wrist-worn devices compared with arm-worn devices [24]. Moreover, arm and wrist movements

cause displacement of the PPG sensor over the skin, alter skin deformation, and affect blood flow dynamics, generating motion artifacts that are difficult to mitigate through filtering or algorithms when occurring frequently and result in false calculations [22,25]. Although the Vantage V2 also uses PPG technology, like the Verity Sense, the difference in wearing position has a great impact on the HR signal quality, requiring distinct filtering methods and algorithms. Similarly, since wrist-worn devices measure at a more distal position, blood flow may be further reduced in cold environments due to vasoconstriction, which has a greater impact on smaller capillaries in the extremities than in the upper arm. Moreover, a good fit on the wrist plays a crucial role in minimizing device movement on the skin, which in turn reduces skin deformation.

In this study, the Vantage V2 performed best during cycling on an ergometer, contrary to the expectation that wrist posture during cycling might negatively impact accuracy [19]. This improved performance could be attributed to ensuring a proper fit of the watch, with the device positioned correctly above the wrist and snugly fitted, which might mitigate issues caused by wrist bending.

Notably, the Vantage V2 showed similar results to, or even outperformed, other wrist-worn devices evaluated in previous studies, particularly during higher-intensity activities. When compared with similar current devices, such as the Garmin Forerunner 945 and Polar Ignite, the Vantage V2 demonstrated slightly higher or similar mean absolute error and MAPE values but exhibited comparable LoAs and slightly stronger positive correlations [54]. In low-intensity activities such as walking, the Vantage V2 showed lower accuracy (ie, higher MAPEs) than the Polar Vantage M and the Garmin Instinct. However, during higher-intensity activities such as jogging and skipping (comparable with HIIT), the Vantage V2 outperformed both devices [28]. During lying, sitting, walking, and squat training (which can be compared with weight training in this study), the Vantage V2 exhibited higher MAPEs in lying and walking but lower MAPEs in sitting and weight training compared with the Fitbit Charge 4 and Samsung Galaxy Watch Active2 [43]. Similarly, in terms of agreement (Pearson correlation), the Vantage V2 exhibited lower agreement in low-intensity activities but outperformed the Apple Watch Series 4, the Polar Vantage V, the Garmin Fenix 5, and the Fitbit Versa at higher HRs [33]. A comparable trend was observed when comparing the Vantage V2 with the Garmin Fenix 6 and the Polar Grit X across various moderate to vigorous activities (eg, walking, incremental maximal treadmill walking, and cycling) [48]. Furthermore, during cycling and resistance training, the Vantage V2 outperformed both the Apple Watch Series 2 and the Bose SoundSport Pulse [42]. The Vantage V2 also showed similar results to those of another study that tested this device in swimming [32].

These findings suggest that the Vantage V2 performs slightly better than its competitors at higher intensities and elevated mean HR, potentially indicating that the device incorporates a robust motion artifact filtering algorithm. However, it remains susceptible to lower blood flow. In summary, while the Vantage V2 still exhibits the typical limitations of wrist-worn sensors, its accuracy is comparable with—or even exceeds—that of some other wrist-worn devices.

### Strengths, Limitations, and Recommendations

This study has several strengths but also faces certain limitations that warrant consideration. First, while the sample size was relatively small and homogeneous in terms of health, age (mean 27.4, SD 5.8 years), and BMI (18.5 - 30 kg/m²), the study benefited from a large dataset (14,653 data points; mean 4.5, SD 2.1 hours per participant). This extensive data volume strengthens the reliability of the analysis and allows for robust analysis. Future research should complement this approach by including a more diverse population to assess broader applicability. Second, the study protocol included a wide range of activities, from sedentary to vigorous intensity, conducted in seminaturalistic conditions in a gymnasium. However, the indoor environment may not fully replicate real-world conditions, and activities outside this range, such as extreme sports or water-based activities, were not evaluated. Third, while the Polar H10 ECG chest strap is a proven criterion measure for HR measurement during various activities and intensities, especially in free-living conditions, the H10 nevertheless exhibited missing data and artifacts in this study, potentially due to suboptimal sensor-wearing position or fitting, or motion-induced signal interference. To mitigate this, rigorous data cleaning and artifact detection procedures were used, including visual screening and the exclusion of outlier activities from the analysis. However, some artifacts may still have introduced variability into the reference data, potentially influencing the comparison with the tested wearable devices. Future studies should be aware of this limitation and carefully review the reference data as well, as errors or artifacts in the reference measurements could lead to misleading comparisons and affect the validity of the findings. Fourth, while the wearing position and fitting of the devices were standardized to ensure consistency, it might not reflect real-world usage where users may wear devices loosely or incorrectly. Including scenarios with varied placement conditions in future studies could better simulate real-world use. Furthermore, device placement on different limbs or at varying positions on the same limb may introduce variability due to differences in blood flow, which was not addressed in this study. Future research should explore whether placing an additional sensor on the same limb influences blood flow and, consequently, HR measurements. Finally, as wearable technologies continue to evolve, continuous validation across various activities, contexts, and populations will be crucial to ensuring that these devices provide accurate and actionable data for health monitoring and the development of physiological metrics (eg, estimation of core body temperature or energy expenditure).

### Conclusions

This study evaluated the accuracy and reliability of 2 currently available wearable devices across a wide range of activities and different wearing positions. The Polar Verity Sense demonstrated excellent accuracy and reliability across a broad range of physical activities and intensities, particularly when worn on the upper arm. The Polar Vantage V2, worn on the wrist, showed overall moderate accuracy and increased

variability. It also demonstrated the typical limitations of wrist-worn devices, including reduced accuracy at lower HRs in combination with arm and wrist movements. However, it demonstrated improved performance at higher intensities and remains a competitive option within its category. These findings highlight the challenges associated with wrist-worn HR devices and the importance of device-wearing position to ensure accurate HR measurements.

In summary, for users seeking valid and reliable HR monitoring across various activities, the Verity Sense presents a strong alternative to ECG-based chest straps. For practical implementation, device selection should be guided by the intended use case, required accuracy, and user needs. Optimizing the chosen device and wearing position is essential to ensuring the highest possible accuracy within its specific context.

## Acknowledgments

The authors would like to thank Alexandre Giroud and Vincent Baeriswyl for their valuable contributions during the data collection.

## Data Availability

The datasets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

TS and RG were involved in the conceptualization. TS was responsible for project management, data collection, statistical analysis, data interpretation, and writing and revising the manuscript. RG was responsible for the project supervision and manuscript revision. Both authors read and approved the final manuscript.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Accuracy and reliability results of the arm-worn Verity Sense (upper arm and forearm).
[DOCX File, 33 KB - cardio_v9i1e67110_app1.docx ]

Multimedia Appendix 2
Accuracy and reliability results of the wrist-worn Polar Vantage V2 (dominant and nondominant wrists).
[DOCX File, 33 KB - cardio_v9i1e67110_app2.docx ]

## References

1. Alugubelli N, Abuissa H, Roka A. Wearable devices for remote monitoring of heart rate and heart rate variability—what we know and what is coming. Sensors (Basel) 2022 Nov 17;22(22):8903. [doi: 10.3390/s22228903] [Medline: 36433498]
2. Zhang D, Wang W, Li F. Association between resting heart rate and coronary artery disease, stroke, sudden death and noncardiovascular diseases: a meta-analysis. CMAJ 2016 Oct 18;188(15):E384-E392. [doi: 10.1503/cmaj.160050] [Medline: 27551034]
3. Adão Martins NR, Annaheim S, Spengler CM, Rossi RM. Fatigue monitoring through wearables: a state-of-the-art review. Front Physiol 2021;12:790292. [doi: 10.3389/fphys.2021.790292] [Medline: 34975541]
4. Buller MJ, Tharion WJ, Cheuvront SN, et al. Estimation of human core temperature from sequential heart rate observations. Physiol Meas 2013 Jul;34(7):781-798. [doi: 10.1088/0967-3334/34/7/781] [Medline: 23780514]
5. Buller MJ, Atkinson E, Driver K, et al. Individualized monitoring of heat illness risk: novel adaptive physiological strain index to assess exercise-heat strain from athletes to fully encapsulated workers. Physiol Meas 2023 Oct 20;44(10):10. [doi: 10.1088/1361-6579/acf991] [Medline: 37703905]
6. Loro FL, Martins R, Ferreira JB, et al. Validation of a wearable sensor prototype for measuring heart rate to prescribe physical activity: cross-sectional exploratory study. JMIR Biomed Eng 2024 Dec 11;9:e57373. [doi: 10.2196/57373] [Medline: 39661434]
7. Hinde K, White G, Armstrong N. Wearable devices suitable for monitoring twenty four hour heart rate variability in military populations. Sensors (Basel) 2021 Feb 4;21(4):1061. [doi: 10.3390/s21041061] [Medline: 33557190]
8. Schaffarczyk M, Rogers B, Reer R, Gronwald T. Validity of the Polar H10 sensor for heart rate variability analysis during resting state and incremental exercise in recreational men and women. Sensors (Basel) 2022 Aug 30;22(17):6536. [doi: 10.3390/s22176536] [Medline: 36081005]
9. Skála T, Vícha M, Rada M, Vácha J, Flašík J, Táborský M. Feasibility of evaluation of Polar H10 chest-belt ECG in patients with a broad range of heart conditions. Cor Vasa 2022 Sep 1;64(4):411-422. [doi: 10.33678/cor.2022.083]

10. Vermunicht P, Makayed K, Meysman P, et al. Validation of Polar H10 chest strap and Fitbit Inspire 2 tracker for measuring continuous heart rate in cardiac patients: impact of artefact removal algorithm. Europace 2023 May 24;25(Suppl 1):euad122.550. [doi: 10.1093/europace/euad122.550]
11. Gilgen-Ammann R, Schweizer T, Wyss T. RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. Eur J Appl Physiol 2019 Jul;119(7):1525-1532. [doi: 10.1007/s00421-019-04142-5] [Medline: 31004219]
12. Beeler N, Roos L, Delves SK, et al. The wearing comfort and acceptability of ambulatory physical activity monitoring devices in soldiers. IISE Trans Occup Ergon Hum Factors 2018 Jan 2;6(1):1-10. [doi: 10.1080/24725838.2018.1435431]
13. Peake JM, Kerr G, Sullivan JP. A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. Front Physiol 2018;9:743. [doi: 10.3389/fphys.2018.00743] [Medline: 30002629]
14. Challoner AV, Ramsay CA. A photoelectric plethysmograph for the measurement of cutaneous blood flow. Phys Med Biol 1974 May;19(3):317-328. [doi: 10.1088/0031-9155/19/3/003] [Medline: 4445210]
15. Düking P, Fuss FK, Holmberg HC, Sperlich B. Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity. JMIR Mhealth Uhealth 2018 Apr 30;6(4):e102. [doi: 10.2196/mhealth.9341] [Medline: 29712629]
16. Witt D, Kellogg R, Snyder M, Dunn J. Windows into human health through wearables data analytics. Curr Opin Biomed Eng 2019 Mar;9:28-46. [doi: 10.1016/j.cobme.2019.01.001] [Medline: 31832566]
17. Rebelo A, Martinho DV, Valente-Dos-Santos J, Coelho-E-Silva MJ, Teixeira DS. From data to action: a scoping review of wearable technologies and biomechanical assessments informing injury prevention strategies in sport. BMC Sports Sci Med Rehabil 2023 Dec 14;15(1):169. [doi: 10.1186/s13102-023-00783-4] [Medline: 38098071]
18. Singh B, Chastin S, Miatke A, et al. Real-world accuracy of wearable activity trackers for detecting medical conditions: systematic review and meta-analysis. JMIR Mhealth Uhealth 2024 Aug 30;12:e56972. [doi: 10.2196/56972] [Medline: 39213525]
19. Barrios L, Oldrati P, Santini S, Lutterotti A. Evaluating the accuracy of heart rate sensors based on photoplethysmography for in-the-wild analysis. Presented at: PervasiveHealth'19: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 20-23, 2019; Trento, Italy p. 251-261. [doi: 10.1145/3329189.3329215]
20. Navalta JW, Montes J, Bodell NG, Salatto RW, Manning JW, DeBeliso M. Concurrent heart rate validity of wearable technology devices during trail running. PLoS One 2020;15(8):e0238569. [doi: 10.1371/journal.pone.0238569] [Medline: 32866216]
21. Navalta JW, Davis DW, Malek EM, et al. Heart rate processing algorithms and exercise duration on reliability and validity decisions in biceps-worn Polar Verity Sense and OH1 wearables. Sci Rep 2023 Jul 20;13(1):11736. [doi: 10.1038/s41598-023-38329-w] [Medline: 37474743]
22. Castaneda D, Esparza A, Ghamari M, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. Int J Biosens Bioelectron 2018;4(4):195-202. [doi: 10.15406/ijbsbe.2018.04.00125] [Medline: 30906922]
23. Shimazaki T, Kuwahara Y, Kimoto M, Hara S, Yomo H. Effect of position and fastening belt pressure on the accuracy of PPG-based heart rate sensor. Presented at: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Jul 18-21, 2018; Honolulu, HI p. 4323-4326. [doi: 10.1109/EMBC.2018.8513333]
24. Tamura T, Maeda Y, Sekine M, Yoshida M. Wearable photoplethysmographic sensors—past and present. Electronics (Basel) 2014;3(2):282-302. [doi: 10.3390/electronics3020282]
25. Zhang Y, Song S, Vullings R, et al. Motion artifact reduction for wrist-worn photoplethysmograph sensors based on different wavelengths. Sensors (Basel) 2019;19(3):673. [doi: 10.3390/s19030673]
26. Mühlen JM, Stang J, Lykke Skovgaard E, et al. Recommendations for determining the validity of consumer wearable heart rate devices: expert statement and checklist of the INTERLIVE Network. Br J Sports Med 2021 Jul;55(14):767-779. [doi: 10.1136/bjsports-2020-103148] [Medline: 33397674]
27. Navalta JW, Carrier B, Blank M, et al. Validity and reliability of wearable technology devices during simulated pickleball game play. Sports (Basel) 2024 Aug 28;12(9):234. [doi: 10.3390/sports12090234] [Medline: 39330711]
28. Navalta JW, Davis DW, Carrier B, et al. Validity and reliability of wearable devices during self-paced walking, jogging and overground skipping. Sport Mont 2023 Oct 1;21(3):23-29. [doi: 10.26773/smj.231004]
29. López-Belmonte Ó, Febles-Castro A, Gay A, Cuenca-Fernández F, Arellano R, Ruiz-Navarro JJ. Validity of the polar verity sense during swimming at different locations and intensities. Scand J Med Sci Sports 2023 Dec;33(12):2623-2625. [doi: 10.1111/sms.14494] [Medline: 37727999]
30. Fullmer WB, Carrier B, Gil D, et al. Validity of average heart rate and energy expenditure in Polar Armband devices while self-paced biking. Med Sci Sports Exerc 2022;54(9S):195-195. [doi: 10.1249/01.mss.0000877508.29887.33]
31. Neudorfer M, Kumar D, Smeddinck JD, et al. Validity of four consumer-grade optical heart rate sensors for assessing volume and intensity distribution of physical activity. Scand J Med Sci Sports 2024 Nov;34(11):e14756. [doi: 10.1111/sms.14756] [Medline: 39508366]

32. Cosoli G, Antognoli L, Veroli V, Scalise L. Accuracy and precision of wearable devices for real-time monitoring of swimming athletes. Sensors (Basel) 2022 Jun 23;22(13):4726. [doi: 10.3390/s22134726] [Medline: 35808223]
33. Düking P, Giessing L, Frenkel MO, Koehler K, Holmberg HC, Sperlich B. Wrist-worn wearables for monitoring heart rate and energy expenditure while sitting or performing light-to-vigorous physical activity: validation study. JMIR Mhealth Uhealth 2020 May 6;8(5):e16716. [doi: 10.2196/16716] [Medline: 32374274]
34. World Health Organization. Physical activity. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2025-02-27]
35. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol 1988 Jun;124(6):869-871. [doi: 10.1001/archderm.124.6.869] [Medline: 3377516]
36. Borg G. Perceived exertion as an indicator of somatic stress. Scand J Rehabil Med 1970;2(2):92-98. [doi: 10.2340/1650197719702239298]
37. Borg E, Borg G. A comparison of AME and CR100 for scaling perceived exertion. Acta Psychol (Amst) 2002 Feb;109(2):157-175. [doi: 10.1016/s0001-6918(01)00055-5] [Medline: 11820425]
38. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020;3:18. [doi: 10.1038/s41746-020-0226-6] [Medline: 32047863]
39. Vinatzer H, Rzepka A, Hayn D, et al. Accuracy of wearable photoplethysmography sensors for continuous heart rate monitoring in telehealth applications. Stud Health Technol Inform 2022 May 16;293:205-211. [doi: 10.3233/SHTI220370] [Medline: 35592983]
40. Goldberger A, Goldberger Z, Shvilkin A. Chapter 13: Sinus and escape rhythms. In: Goldberger's Clinical Electrocardiography, 8th edition: WB Saunders; 2013:114-120. [doi: 10.1016/B978-0-323-08786-5.00013-0]
41. Joosen P, Exadaktylos V, Taelman J, Aerts JM, Berckmans D. Method for artefact detection and removal in heart rate signals measured during physical exercise. 2013 Presented at: Proceedings of the International Congress on Cardiovascular Technologies; Villamoura, Portugal p. 57-61. [doi: 10.5220/0004636800570061]
42. Boudreaux BD, Hebert EP, Hollander DB, et al. Validity of wearable activity monitors during cycling and resistance exercise. Med Sci Sports Exerc 2018 Mar;50(3):624-633. [doi: 10.1249/MSS.0000000000001471] [Medline: 29189666]
43. Nissen M, Slim S, Jäger K, et al. Heart rate measurement accuracy of Fitbit Charge 4 and Samsung Galaxy Watch Active2: device evaluation study. JMIR Form Res 2022 Mar 1;6(3):e33635. [doi: 10.2196/33635] [Medline: 35230250]
44. McBride GB. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. NIWA client report: HAM2005-062. : National Institute of Water & Atmospheric Research Ltd; 2005 URL: https://www.medcalc.org/download/pdf/McBride2005.pdf [accessed 2024-12-21]
45. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989 Mar;45(1):255-268. [doi: 10.2307/2532051] [Medline: 2720055]
46. Steichen TJ, Cox NJ. A note on the concordance correlation coefficient. Stata J Promot Commun Stat Stata 2002 Jun;2(2):183-189. [doi: 10.1177/1536867X0200200206]
47. Hopkins W. Validity thresholds and error rates for test measures used to assess individuals. In: Baca A, Wessner B, Diketmüller R, Tschan H, Hofmann M, Kornfeind P, et al, editors. Book of Abstracts of the 21st Annual Congress of the European College of Sport Science: European College of Sport Science; 2016:6-9.
48. Merrigan JJ, Stovall JH, Stone JD, Stephenson M, Finomore VS, Hagen JA. Validation of Garmin and Polar devices for continuous heart rate monitoring during common training movements in tactical populations. Meas Phys Educ Exerc Sci 2023 Jul 3;27(3):234-247. [doi: 10.1080/1091367X.2022.2161820]
49. Pasadyn SR, Soudan M, Gillinov M, et al. Accuracy of commercially available heart rate monitors in athletes: a prospective study. Cardiovasc Diagn Ther 2019 Aug;9(4):379-385. [doi: 10.21037/cdt.2019.06.05]
50. Thomson EA, Nuss K, Comstock A, et al. Heart rate measures from the Apple Watch, Fitbit Charge HR 2, and electrocardiogram across different exercise intensities. J Sports Sci 2019 Jun;37(12):1411-1419. [doi: 10.1080/02640414.2018.1560644] [Medline: 30657025]
51. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg 2018 May;126(5):1763-1768. [doi: 10.1213/ANE.0000000000002864] [Medline: 29481436]
52. Chow HW, Yang CC. Accuracy of optical heart rate sensing technology in wearable fitness trackers for young and older adults: validation and comparison study. JMIR Mhealth Uhealth 2020 Apr 28;8(4):e14707. [doi: 10.2196/14707] [Medline: 32343255]
53. Müller AM, Wang NX, Yao J, et al. Heart rate measures from wrist-worn activity trackers in a laboratory and free-living setting: validation study. JMIR Mhealth Uhealth 2019 Oct 2;7(10):e14120. [doi: 10.2196/14120] [Medline: 31579026]
54. Budig M, Keiner M, Stoohs R, Hoffmeister M, Höltke V. Heart rate and distance measurement of two multisport activity trackers and a cell phone app in different sports: a cross-sectional validation and comparison field study. Sensors (Basel) 2021 Dec 28;22(1):180. [doi: 10.3390/s22010180] [Medline: 35009723]


## Abbreviations

**CCC:** concordance correlation coefficient

**ECG:** electrocardiogram
**HIIT:** high-intensity interval training
**HR:** heart rate
**LoA:** limits of agreement
**MAE:** mean absolute error
**MAPE:** mean absolute percentage error
**PPG:** photoplethysmography
**RMSE:** root-mean-square error
**WSCV:** within-subject coefficient of variation

*Edited by A Coristine; submitted 02.10.24; peer-reviewed by J Navalta, JM Aerts; revised version received 11.02.25; accepted 11.02.25; published 21.03.25.*

*Please cite as:*
*Schweizer T, Gilgen-Ammann R*
*Wrist-Worn and Arm-Worn Wearables for Monitoring Heart Rate During Sedentary and Light-to-Vigorous Physical Activities: Device Validation Study*
*JMIR Cardio 2025;9:e67110*
*URL: https://cardio.jmir.org/2025/1/e67110*
*doi:10.2196/67110*

# The Effect of Discharge Planning Videos and Booklets on Quality of Life Among Patients With Heart Failure: Quasi-Experimental Study

Fitri Arofiati[1], PhD; Fuji Dwi Lestari[1], MNS; Henri Setiawan[2], PhD

[1]Department of Nursing, Universitas Muhammadiyah Yogyakarta, Jl. Brawijaya, Kasihan, Bantul, Indonesia

[2]Department of Nursing, STIKes Muhammadiyah Ciamis, Ciamis, Indonesia

**Corresponding Author:**
Fitri Arofiati, PhD
Department of Nursing, Universitas Muhammadiyah Yogyakarta, Jl. Brawijaya, Kasihan, Bantul, Indonesia

## Abstract

**Background:** Heart failure remains a major global health issue, significantly impacting patients' quality of life due to its chronic and progressive nature. Effective discharge planning, including educational interventions such as videos and booklets, plays a crucial role in enhancing self-care management and overall patient well-being.

**Objective:** The aim of this study is to evaluate the effects of discharge planning videos and booklets on the quality of life of patients with heart failure.

**Methods:** This study used a quasi-experimental design and was conducted at PKU Muhammadiyah Gamping Hospital from July to November 2024. A total of 42 participants who met the inclusion criteria were selected based on sample size calculations using G*Power and were evenly assigned to intervention and control groups. Both groups received standard discharge planning provided by health care professionals. Discharge planning videos and booklets were developed as educational tools for the intervention group. The Minnesota Living With Heart Failure Questionnaire was used to assess quality of life. The independent sample *t* test was used to analyze the effect of the intervention using SPSS (version 29). This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and was approved by the institutional review board (number 150/KEP-PKU/VII/2024).

**Results:** The intervention significantly improved the quality of life of patients with heart failure, with the mean score decreasing from 39.00 (SD 8.11) to 24.76 (SD 4.02; $P<.001$) in the intervention group. In contrast, the control group showed minimal change, from 39.90 (SD 5.89) to 40.24 (SD 5.84), resulting in a statistically significant between-group difference of 15.58 ($P<.001$). Furthermore, the effect size was large (Cohen $d=3.09$), suggesting a strong practical significance of the intervention in enhancing the quality of life among patients with heart failure. Moreover**,** the mean Minnesota Living With Heart Failure Questionnaire scores across 4 domains—physical, mental, emotional, and social—also showed significant improvements after the intervention. The intervention group experienced reductions in all domains: physical (9.95 to 6.76), mental (7.81 to 5.62), emotional (13.19 to 7.48), and social (8.05 to 4.90), whereas the control group showed minimal or no change. These results indicate that the intervention effectively improved patients' quality of life across multiple dimensions.

**Conclusions:** Discharge planning through videos and booklets may improve the quality of life of patients with heart failure compared to standard care. These findings highlight the potential clinical value of structured patient education. The intervention appeared to enhance patients' understanding of their condition and support self-management behaviors, including adherence to lifestyle recommendations. However, they should be interpreted with caution and confirmed through further studies with larger and more diverse populations.

***(JMIR Cardio 2025;9:e75417)*** doi:10.2196/75417

**KEYWORDS**

discharge planning; heart failure; patient-centered care; quality of life; video

## Introduction

Heart failure is a chronic and progressive condition that significantly impairs the quality of life (QoL) of patients globally. Characterized by the heart's inability to pump blood efficiently, it manifests through symptoms such as shortness of breath, fatigue, and fluid retention. Despite advancements in medical treatments, heart failure remains a leading cause of hospitalization and mortality, imposing a substantial burden on health care systems and patients alike [1-3]. In Indonesia, heart failure is among the top 10 noncommunicable diseases, with 229,696 (0.13%) patients diagnosed with the condition.

Additionally, based on doctor's diagnoses or symptoms, the prevalence of heart failure in Indonesia is estimated at 1.5%, affecting approximately 2,650,340 individuals [4]. These alarming statistics highlight the urgent need for effective management strategies to mitigate the growing burden of heart failure.

The management of heart failure is multifaceted, requiring not only pharmacological interventions but also comprehensive patient education and self-care strategies [5,6]. Unhealthy lifestyles and the inability of patients to independently manage their condition are significant contributors to the rising incidence of heart failure. Addressing these factors necessitates a focus on enhancing self-management among patients with heart failure. In this context, discharge planning emerges as a critical component in ensuring that patients are equipped with the necessary knowledge and skills to manage their condition effectively after leaving the hospital [7,8].

Discharge planning is a multidisciplinary process designed to facilitate the transition of patients from hospital to home, thereby reducing the risk of readmission and improving overall health outcomes. It involves the coordination of care, patient education, and the provision of resources to support self-management [7]. Numerous studies have underscored the importance of discharge planning in improving health outcomes for patients with heart failure. For instance, Rice et al [9] demonstrated that comprehensive discharge planning and postdischarge support significantly enhance health outcomes, potentially improving patients' QoL through better education and resources [9]. Similarly, Graupner et al [10] found that structured discharge planning interventions led to improved outcomes, including enhanced QoL, increased knowledge related to heart failure, improved self-care behaviors, and reduced readmission rates. These findings underscore the critical role of well-designed discharge planning programs in addressing the multifaceted challenges faced by patients with heart failure.

Patient education is a cornerstone of effective heart failure management. Empowering patients with the knowledge and skills to actively participate in their care can lead to better adherence to treatment regimens, improved symptom management, and reduced hospital readmissions [11-13]. However, many patients face challenges such as low health literacy, cognitive impairments, or language barriers, which can hinder their ability to understand and apply the information provided. Additionally, the emotional and psychological burden of living with a chronic condition like heart failure can further complicate the educational process [14,15].

In recent years, the advent of digital technology has opened new avenues for enhancing patient education and engagement [16]. Multimedia tools, such as videos, offer a promising solution by presenting information in a clear, concise, and visually appealing manner. Videos can be tailored to address the specific needs and preferences of individual patients, making the content more relevant and effective [17-19]. For example, videos can demonstrate proper techniques for monitoring blood pressure, taking medications, or performing physical exercises, providing patients with practical guidance that they can easily follow at home. Moreover, videos can be accessed repeatedly, allowing patients to review the information as needed, which reinforces learning and promotes long-term retention [20-22].

Although videos offer a dynamic and engaging medium for patient education, booklets remain a valuable complementary tool. Booklets provide a written reference that patients can consult at their own pace, offering detailed information on various aspects of heart failure management, such as dietary recommendations, medication schedules, and warning signs of worsening symptoms [23,24]. They can also include diagrams, charts, and checklists to facilitate their understanding and application of the information. Furthermore, booklets can be customized to reflect the cultural and linguistic diversity of the patient population, ensuring that the content is accessible and relevant to all [25-27]. The integration of videos and booklets in discharge planning may offer a synergistic effect, enhancing the overall quality of patient education. Videos can capture the patient's attention and convey key messages in an engaging manner, while booklets provide a comprehensive resource for additional details and clarification [28,29]. Together, these tools address the cognitive, emotional, and practical aspects of patient education, promoting a more holistic approach to self-care. Therefore, this study aims to evaluate the effects of discharge planning videos and booklets on the QoL of patients with heart failure. It is hypothesized that patients who receive discharge planning with video and booklet support will experience a significantly greater improvement in QoL compared to those who receive standard discharge planning alone.

## Methods

### Design and Setting

This quasi-experimental study was conducted at PKU Muhammadiyah Gamping Hospital from July to November 2024. Such studies offer a valuable alternative for estimating causal relationships and are increasingly used as more observational data become available [30].

### Participants and Sampling

The study population consisted of patients with heart failure hospitalized at PKU Muhammadiyah Gamping Hospital. The sample size was calculated using G*Power (version 3.1), applying a *t* test for independent means (two groups) with a significance level of 0.05, statistical power of 0.80, and an effect size of 0.80 (large). A large effect size (Cohen *d*=0.8) was assumed based on prior studies demonstrating marked improvements in the QoL of patients with heart failure following structured educational interventions [31,32]. A total of 42 respondents were recruited and equally divided into two groups: 21 in the intervention group and 21 in the control group. Inclusion criteria included patients who had heart failure classified as grade 1 or 2, were aged over 20 years, were literate, and were smartphone users. Exclusion criteria included patients who died, were readmitted within a month, or had more than 3 comorbidities.

### Intervention

Both groups received standard discharge planning provided by health care professionals. In addition, the intervention group received supplementary educational materials in the form of a

video and a booklet, specifically developed to enhance patients' understanding of self-care following hospital discharge. The content of these materials was guided by the Self-Care of Heart Failure Index framework and emphasized 3 core components: symptom monitoring, adherence to treatment, and self-care management.

The educational video, lasting 4 minutes and 27 seconds, included information on the definition and symptoms of heart failure, contributing risk factors, dietary recommendations, medication adherence, and strategies for home-based care. The video was uploaded to YouTube to ensure ease of access and could be rewatched as needed by patients or their caregivers [33]. The accompanying booklet (Multimedia Appendix 1) served as a written reference that reinforced the video content and included illustrations and simple language tailored to patients with varying literacy levels.

Both the video and booklet were developed collaboratively by a team of nursing lecturers, cardiologists, and cardiology ward nurses to ensure clinical accuracy and contextual relevance. The materials were reviewed through an expert validation process involving 2 cardiologists and 3 senior nurses using a structured content validity checklist. Additionally, the materials were pilot-tested with a group of 5 patients with heart failure to assess clarity, usability, and acceptability. Feedback from the pilot testing was used to refine the wording, visuals, and delivery method of the materials before full implementation in the study.

### Outcome Measurement: QoL

The Minnesota Living With Heart Failure Questionnaire (MLHFQ) was used to assess QoL across 4 domains: physical, emotional, mental, and social. The Indonesian version of the MLHFQ consists of 20 validated items (excluding question 10 due to low item correlation), scored using a 4-point Likert scale. Lower scores indicate better QoL, whereas higher scores reflect poorer perceived health status. Total scores range from 24 to 80, categorized as <24 (good), 24 - 45 (moderate), and >45 (poor). The instrument showed high reliability (Cronbach $\alpha$=0.954) [34,35].

### Data Collection and Analysis

Data collection occurred in 2 phases: a pretest before intervention and a posttest 4 weeks after intervention. Patients completed the MLHFQ questionnaire at both time points to assess changes in QoL (Multimedia Appendix 2). All data were analyzed using SPSS (version 29; IBM Corp). Descriptive statistical analysis was conducted for sociodemographic variables, including age, sex, education, and occupation. These characteristics were presented using frequencies and percentages for each group. The Shapiro-Wilk test was also used to assess the homogeneity of sociodemographic data distributions between groups. The Shapiro-Wilk test confirmed normal distribution (Multimedia Appendix 3), allowing for parametric tests. Paired *t* tests were used to compare pretest and posttest scores within groups, while independent *t* tests compared differences between groups at each time point.

### Ethical Considerations

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and was approved by the institutional review board of PKU Muhammadiyah Gamping Hospital, Indonesia (150/KEP-PKU/VII/2024). All participants provided written informed consent prior to participation, which included consent for the use of their data in secondary analyses. The confidentiality and privacy of participants were protected by using anonymized codes in all data records, and no personally identifiable information was collected or reported. Participants did not receive any monetary or material compensation for their involvement in the study. Additionally, no images or materials containing identifiable features of individual participants are included in this manuscript or its supplementary files.

## Results

### Sociodemographics of the Participants

Detailed information on participant flow and allocation can be seen in Figure 1. Table 1 summarizes the sociodemographic characteristics of participants (n=42) in both groups. The mean age was comparable (intervention: 62.8, SD 12.3; control: 60.7, SD 15.7 y), with a higher proportion of males in both groups. Primary education was most common, and more participants in the intervention group were employed. All *P* values were >.05, indicating no significant baseline differences between groups, suggesting balanced sociodemographic characteristics.

**Figure 1.** Flow diagram of participant attrition. ITT: intention-to-treat.

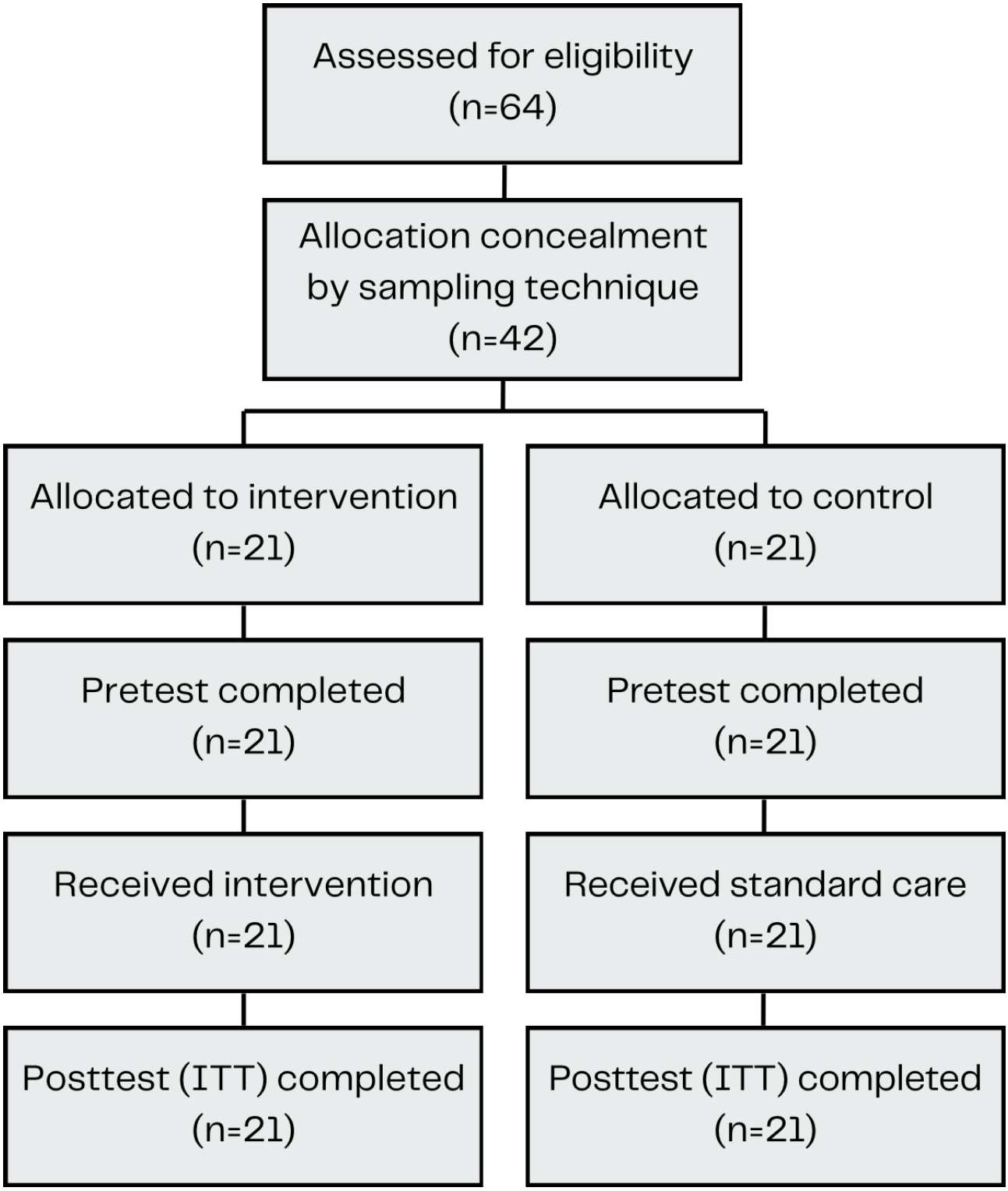


**Table .** Sociodemographics of the participants.

| Variables | | Group | | *P* value[a] |
|---|---|---|---|---|
| | | Intervention | Control | |
| Age, mean (SD) | | 62.8 (12.3) | 60.7 (15.7) | .28 |
| Age, n (%) | | | | |
| | <Mean | 9 (42.9) | 10 (47.6) | |
| | >Mean | 12 (57.1) | 11 (52.4) | |
| Sex, n (%) | | | | .84 |
| | Male | 15 (71.4) | 12 (57.1) | |
| | Female | 6 (28.6) | 9 (42.9) | |
| Education, n (%) | | | | .67 |
| | Primary | 5 (23.8) | 10 (47.6) | |
| | Secondary | 2 (9.5) | 2 (9.5) | |
| | Tertiary | 10 (47.6) | 8 (38.1) | |
| | University | 4 (19) | 1 (4.8) | |
| Occupation, n (%) | | | | .31 |
| | Employed | 13 (61.9) | 10 (47.6) | |
| | Unemployed | 8 (38.1) | 11 (52.4) | |

[a]Homogenity test by Shapiro-Wilk.

## The Effect of the Intervention on the QoL of Patients With Heart Failure

Table 2 presents the impact of discharge planning using videos and booklets on the QoL of patients with heart failure. Before the intervention, the mean scores were similar between groups (intervention: 39.00, SD 8.11; control: 39.90, SD 5.89). After the intervention, the intervention group's mean score significantly decreased to 24.76 (SD 4.02; *P*<.001), while the control group showed minimal change (40.24, SD 5.84; *P*=.031). The postintervention difference between groups was statistically significant (mean difference=15.58; *P*<.001), indicating that the intervention substantially improved patients' QoL compared to standard discharge planning. Furthermore, the effect size was large (Cohen *d*=3.09), suggesting a strong practical significance of the intervention in enhancing QoL among patients with heart failure.

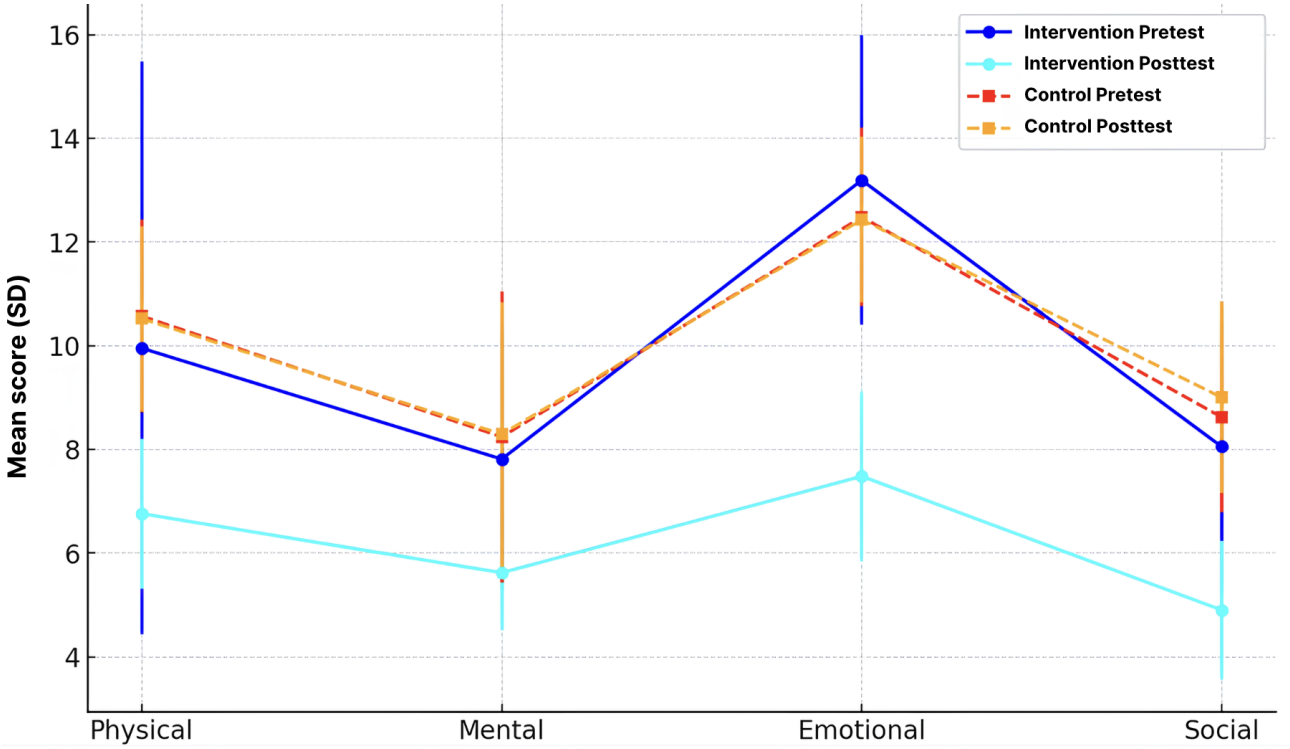
**Table .** The effect of the intervention on the quality of life of patients with heart failure.

| Parameters | Group | |
|---|---|---|
| | Intervention | Control |
| Pretest, mean (SD) | 39.00 (8.11) | 39.90 (5.89) |
| Posttest[a], mean (SD) | 24.76 (4.02) | 40.24 (5.84) |
| Significance[b] | <.001 | .03 |

[a]Posttest, the mean difference between the intervention and control group was 15.58 (independent sample *t* test *P*<.001; Cohen *d*=3.09).
[b]Paired *t* test.

Figure 2 illustrates the mean MLHFQ scores across 4 domains—physical, mental, emotional, and social—before and after the intervention. The intervention group showed significant reductions in all domains: physical (9.95 to 6.76), mental (7.81 to 5.62), emotional (13.19 to 7.48), and social (8.05 to 4.90). In contrast, the control group showed minimal or no change. These results indicate that the intervention effectively improved patients' QoL across multiple dimensions.

**Figure 2.** Comparison of mean scores and standard deviation between two groups.




## Discussion

### Principal Findings

The results of this study demonstrated a statistically significant improvement in the QoL of patients with heart failure who received discharge planning interventions involving educational videos and booklets, as evidenced by the independent sample *t* test ($P<.001$). Although the control group exhibited a statistically significant change, the numerical difference was minimal and is unlikely to represent a clinically meaningful effect. This change may reflect random sampling variability rather than a true treatment-related outcome. This significant difference underscores the effectiveness of these multimedia tools in enhancing patients' understanding of self-care, promoting adherence to management strategies, and ultimately improving their overall QoL compared to the control group. The findings suggest that integrating videos and booklets into discharge planning can serve as a valuable approach in clinical settings, offering a practical and accessible means to empower patients and address the multifaceted challenges of heart failure management [36,37]. These results align with previous research emphasizing the importance of structured patient education and support in improving health outcomes, further validating the potential of such interventions to reduce the burden of heart failure and enhance patients' well-being [38,39].

### Multimedia Tools in Patient Education

The use of educational videos in discharge planning likely contributed to the observed improvements by presenting complex medical information in a clear, engaging, and visually appealing manner. Murphy et al [40] emphasized that the dynamic nature of videos allows for the inclusion of visual and auditory elements, which cater to different learning styles, making the information more accessible to a broader audience. Similarly, Saluky and Bahiyah [41] highlighted that the ability to revisit video content allows patients to reinforce comprehension, address areas of difficulty, and improve long-term retention of key information. This aligns with studies highlighting the advantages of multimedia tools in overcoming barriers such as low health literacy, cognitive impairments, and language difficulties, which are common among patients with heart failure. By addressing these barriers, videos can enhance patients' confidence and ability to manage their condition independently, ultimately leading to better health outcomes [42,43].

The inclusion of booklets in the intervention provided a complementary resource that allowed patients to access detailed written information at their convenience. Booklets serve as a reliable reference for patients, offering step-by-step guidance on dietary recommendations, exercise routines, and medication management [44-46]. This dual approach—combining the dynamic nature of videos with the comprehensive detail of booklets—likely created a synergistic effect, enhancing the overall impact of the discharge planning intervention. The combination of these tools addresses both the immediate and long-term educational needs of patients, providing them with the resources necessary to manage their condition effectively over time [47-50].

### Psychosocial Benefits of Multimedia-Based Discharge Planning

The significant improvement in QoL observed in this study also highlights the importance of addressing the emotional and psychological aspects of living with heart failure. Tsabedze et al [51] stated that depression and anxiety symptoms were found in over half of patients attending the congestive heart failure clinic, highlighting how chronic conditions like heart failure often lead to emotional distress and a sense of helplessness,

which can hinder patients' ability to engage in self-care [52]. By providing clear, actionable information through videos and booklets, the intervention may have alleviated some of these emotional burdens, empowering patients to take control of their health. This is particularly important given the strong link between psychological well-being and adherence to treatment regimens. The positive outcomes observed in this study suggest that multimedia-based discharge planning can play a crucial role in fostering a sense of empowerment and resilience among patients with heart failure [53,54].

### Study Implications

The findings of this study have important implications for health care systems, particularly in resource-limited settings. Heart failure is a global health challenge that places a significant burden on health care infrastructure, with high rates of hospitalization and readmission [55,56]. Effective discharge planning interventions, such as the use of videos and booklets, offer a cost-effective and scalable solution to improve patient outcomes and reduce health care costs [57-60]. These tools can be easily disseminated and adapted to meet the needs of diverse patient populations, making them a viable option for widespread implementation. Additionally, the use of multimedia tools aligns with the growing trend of digital health solutions, which have the potential to revolutionize patient education and self-management. By leveraging technology, health care providers can deliver high-quality education to patients in a format that is both accessible and engaging [16,61].

### Strengths and Limitations

One of the key strengths of this study lies in its innovative approach to discharge planning, which combines educational videos and booklets to address the diverse learning needs of patients with heart failure. Additionally, the study adopts a holistic perspective, focusing not only on clinical outcomes but also on the emotional and psychological challenges faced by patients. This patient-centered design empowers individuals with knowledge and practical tools, promoting better self-care and overall QoL. The intervention's scalability and accessibility further strengthen its potential, as videos and booklets are cost-effective and can be adapted to various health care settings, including resource-limited environments.

However, this study has several limitations that should be considered. First, the quasi-experimental design, while practical for real-world settings, lacks the methodological rigor of a randomized controlled trial, which limits the strength of causal inferences. Second, the short follow-up period of 4 weeks restricts the ability to assess the long-term sustainability of improvements in QoL and self-care. Third, although the intervention was designed to be scalable, its implementation in resource-limited settings may face challenges related to limited technological access, funding constraints, and staff capacity. Fourth, the study did not account for several external variables such as socioeconomic status, family support, and comorbidities, which could have influenced patient outcomes and introduced potential confounding. Fifth, although baseline measures between groups were relatively similar, the analysis did not statistically adjust for baseline differences, which may limit the precision of the estimated intervention effects, particularly given the small sample size.

### Recommendations and Future Work

Given the positive outcomes observed, future studies should explore the long-term effects of discharge education using multimedia and printed materials, particularly in diverse health care settings and among patients with varying levels of health literacy. It is also recommended to conduct randomized controlled trials to strengthen causal inferences and examine the cost-effectiveness of such interventions. Additionally, integrating digital tools such as mobile health apps could be considered to further support patient self-management beyond hospital discharge.

### Conclusions

This study suggests that discharge planning incorporating videos and booklets may help improve the QoL of patients with heart failure compared to standard care. The intervention appeared to enhance patients' understanding of their condition and support self-management behaviors, including adherence to lifestyle recommendations. Although the results indicate the potential value of structured, multimedia-based patient education, these findings should be interpreted with caution and considered preliminary. Further research with larger, more diverse populations is recommended to confirm the observed effects and assess broader clinical applicability.

## Acknowledgments

This research was funded by Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi through the Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat under contract number 107/E5/PG.02.00.PL/2024.

## Authors' Contributions

Conceptualization: FA; data curation: FDL; formal analysis: HS; investigation: FA and FDL; supervision: FA; writing the original draft: FA and FDL; writing the review and editing: HS. All authors read and agreed to the final version of the manuscript. All authors have critically reviewed and approved the final draft and are responsible for the content and similarity index of the manuscript.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Booklet of discharge planning.
[PDF File, 9418 KB - cardio_v9i1e75417_app1.pdf ]

Multimedia Appendix 2
Minnesota Living With Heart Failure Questionnaire.
[PDF File, 129 KB - cardio_v9i1e75417_app2.pdf ]

Multimedia Appendix 3
Normality test results.
[DOCX File, 14 KB - cardio_v9i1e75417_app3.docx ]

## References


1. Kularatna S, Jadambaa A, Hewage S, Brain D, McPhail S, Parsonage W. Global, regional, and national burden of heart failure associated with atrial fibrillation. BMC Cardiovasc Disord 2023 Jul 11;23(1):345. [doi: 10.1186/s12872-023-03375-9] [Medline: 37430216]
2. Ferreir JL, Kumar S, Soni A, et al. Clinical management of cardiovascular diseases. In: Current Trends in the Diagnosis and Management of Metabolic Disorders 2023. [doi: 10.1201/9781003384823-15]
3. Elendu C, Amaechi DC, Elendu TC, et al. A comprehensive review of heart failure: unraveling the etiology, decoding pathophysiological mechanisms, navigating diagnostic modalities, exploring pharmacological interventions, advocating lifestyle modifications, and charting the horizon of emerging therapies in the complex landscape of chronic cardiac dysfunction. Medicine (Baltimore) 2024 Jan 19;103(3):e36895. [doi: 10.1097/MD.0000000000036895] [Medline: 38241566]
4. Nursita H, Pratiwi A. Peningkatan Kualitas Hidup Pada Pasien Gagal Jantung: a narrative review article [Article in Indonesian]. Jurnal Berita Ilmu Keperawatan 2020;13(1). [doi: 10.23917/bik.v13i1.11916]
5. Jaarsma T, Hill L, Bayes-Genis A, et al. Self-care of heart failure patients: practical management recommendations from the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2021 Jan;23(1):157-174. [doi: 10.1002/ejhf.2008] [Medline: 32945600]
6. Butler J, Petrie MC, Bains M, et al. Challenges and opportunities for increasing patient involvement in heart failure self-care programs and self-care in the post-hospital discharge period. Res Involv Engagem 2023 Apr 13;9(1):23. [doi: 10.1186/s40900-023-00412-x] [Medline: 37046357]
7. Abouzid MR, Siddiqi M. Discharge planning and resource utilization for minimizing readmissions in acute exacerbation of chronic heart failure: insights from an observational study in a community hospital. Curr Probl Cardiol 2024 Feb;49(2):102197. [doi: 10.1016/j.cpcardiol.2023.102197] [Medline: 37977415]
8. Davisson E, Swanson E. Nurses' heart failure discharge planning part II: implications for the hospital system. Appl Nurs Res 2020 Dec;56:151336. [doi: 10.1016/j.apnr.2020.151336] [Medline: 32741656]
9. Rice H, Say R, Betihavas V. The effect of nurse-led education on hospitalisation, readmission, quality of life and cost in adults with heart failure. A systematic review. Patient Educ Couns 2018 Mar;101(3):363-374. [doi: 10.1016/j.pec.2017.10.002] [Medline: 29102442]
10. Graupner C, Kimman ML, Mul S, et al. Patient outcomes, patient experiences and process indicators associated with the routine use of patient-reported outcome measures (PROMs) in cancer care: a systematic review. Support Care Cancer 2021 Feb;29(2):573-593. [doi: 10.1007/s00520-020-05695-4] [Medline: 32875373]
11. Adinkrah E. Patient education in chronic disease management. Physician Assistant Clinics 2024 Oct;9(4):527-540. [doi: 10.1016/j.cpha.2024.05.005]
12. Unverzagt S, Meyer G, Mittmann S, et al. Improving treatment adherence in heart failure: a systematic review and meta-analysis of pharmacological and lifestyle interventions. Dtsch Arztebl Int 2016;113(25). [doi: 10.3238/arztebl.2016.0423] [Medline: 27397013]
13. Morley CM, Levin SA. Health literacy, health confidence, and simulation: a novel approach to patient education to reduce readmissions. Prof Case Manag 2021;26(3):138-149. [doi: 10.1097/NCM.0000000000000456] [Medline: 33784721]
14. Heijmans M, Waverijn G, Rademakers J, van der Vaart R, Rijken M. Functional, communicative and critical health literacy of chronic disease patients and their importance for self-management. Patient Educ Couns 2015 Jan;98(1):41-48. [doi: 10.1016/j.pec.2014.10.006] [Medline: 25455794]
15. Hersh L, Salzman B, Snyderman D. Health literacy in primary care practice. Am Fam Physician 2015 Jul 15;92(2):118-124. [Medline: 26176370]
16. Setiawan H, Xu W, Wang C, et al. The effect of mobile application based genetic counseling on the psychosocial well-being of thalassemia patients and caregivers: a randomized controlled trial. Patient Educ Couns 2025 Jan;130:108457. [doi: 10.1016/j.pec.2024.108457] [Medline: 39418673]
17. Tom K, Phang PT. Effectiveness of the video medium to supplement preoperative patient education: a systematic review of the literature. Patient Educ Couns 2022 Jul;105(7):1878-1887. [doi: 10.1016/j.pec.2022.01.013] [Medline: 35101307]

18. Vandelanotte C, Short C, Plotnikoff RC, et al. TaylorActive--examining the effectiveness of web-based personally-tailored videos to increase physical activity: a randomised controlled trial protocol. BMC Public Health 2015 Oct 5;15(1):1020. [doi: 10.1186/s12889-015-2363-4] [Medline: 26438225]
19. Khalil C, Van Deen W, Dupuy T, Bonthala N, Almario C, Spiegel B. Developing patient-centered inflammatory bowel disease-related educational videos optimized for social media: qualitative research study. JMIR Med Educ 2020 Oct 20;6(2):e21639. [doi: 10.2196/21639] [Medline: 33079065]
20. McDonough DJ, Helgeson MA, Liu W, Gao Z. Effects of a remote, YouTube-delivered exercise intervention on young adults' physical activity, sedentary behavior, and sleep during the COVID-19 pandemic: randomized controlled trial. J Sport Health Sci 2022 Mar;11(2):145-156. [doi: 10.1016/j.jshs.2021.07.009] [Medline: 34314877]
21. Fiallos A, Fiallos C, Figueroa S. TikTok and education: discovering knowledge through learning videos. Presented at: 2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG); Jul 28-30, 2021; Quito, Ecuador. [doi: 10.1109/ICEDEG52154.2021.9530988]
22. McManus RJ, Little P, Stuart B, et al. HOME and Online Management and Evaluation of Blood Pressure (HOME BP) using a digital intervention in poorly controlled hypertension: randomised controlled trial. BMJ 2021 Jan 19;372:m4858. [doi: 10.1136/bmj.m4858] [Medline: 33468518]
23. Morrison AK, Glick A, Yin HS. Health literacy: implications for child health. Pediatr Rev 2019 Jun;40(6):263-277. [doi: 10.1542/pir.2018-0027] [Medline: 31152099]
24. Montagnese C, Santarpia L, Iavarone F, et al. Food-based dietary guidelines around the world: Eastern Mediterranean and Middle Eastern countries. Nutrients 2019 Jun 13;11(6):1325. [doi: 10.3390/nu11061325] [Medline: 31200446]
25. Keinki C, Rudolph I, Ruetters D, et al. The quality of patient information booklets for cancer patients-an evaluation of free accessible material in German language. J Cancer Educ 2018 Oct;33(5):1132-1139. [doi: 10.1007/s13187-017-1224-5] [Medline: 28474220]
26. King-Shier K, Lau A, Fung S, LeBlanc P, Johal S. Ethnocultural influences in how people prefer to obtain and receive health information. J Clin Nurs 2018 Apr;27(7-8):e1519-e1528. [doi: 10.1111/jocn.14281] [Medline: 29396869]
27. Lambert S, Loban E, Li J, et al. Chronic illness management in culturally and linguistically diverse patients: exploring the needs, access, and understanding of information. Qual Health Res 2021 Nov;31(13):2426-2439. [doi: 10.1177/10497323211040769] [Medline: 34636279]
28. Wild CF, Nietsche EA, Salbego C, Teixeira E, Favero NB. Validation of educational booklet: an educational technology in dengue prevention. Rev Bras Enferm 2019 Sep 16;72(5):1318-1325. [doi: 10.1590/0034-7167-2018-0771] [Medline: 31531657]
29. Feeley TH, Keller M, Kayler L. Using animated videos to increase patient knowledge: a meta-analytic review. Health Educ Behav 2023 Apr;50(2):240-249. [doi: 10.1177/10901981221116791] [Medline: 35950326]
30. Carter AW, Jayawardana S, Costa-Font J, Nasir K, Krumholz HM, Mossialos E. How to use quasi-experimental methods in cardiovascular research: a review of current practice. Circ Cardiovasc Qual Outcomes 2024 Feb;17(2):e010078. [doi: 10.1161/CIRCOUTCOMES.123.010078] [Medline: 38362765]
31. Ünalan A. Sample size in clinical researches: power of the test and effect size. Black Sea Journal of Health Science 2021;4(3):221-227. [doi: 10.19127/bshealthscience.866556]
32. Chanuan U, Kajohnsak C, Nittaya S. Sample size estimation using Yamane and Cochran and Krejcie and Morgan and Green formulas and Cohen statistical power analysis by G*Power and comparisons. Apheit International Journal 2021;10(2) [FREE Full text]
33. Vidio Edukasi Kesehatan pada pasien Gagal Jantung. YouTube. URL: https://www.youtube.com/watch?v=oYrOHsPA5ZA [accessed 2025-08-22]
34. Middel B, Bouma J, de Jongste M, et al. Psychometric properties of the Minnesota Living with Heart Failure Questionnaire (MLHF-Q). Clin Rehabil 2001 Oct;15(5):489-500. [doi: 10.1191/026921501680425216] [Medline: 11594639]
35. Rahmani S, Andean R. Correlation between Six-Minutes' Walk Test and quality of life in heart failure patients. amj 2015 Dec;2(4). [doi: 10.15850/amj.v2n4.629]
36. Rad MG, Ghanbari-Afra L, Hoseini MHM, Afra MG, Asayesh H. Effectiveness of self-care program on the quality of life in patients with coronary artery disease undergoing cardiac rehabilitation: a randomized clinical trial. J Educ Health Promot 2021;10:375. [doi: 10.4103/jehp.jehp_70_21] [Medline: 34912911]
37. Deshpande N, Wu M, Kelly C, et al. Video-based educational interventions for patients with chronic illnesses: systematic review. J Med Internet Res 2023 Jul 19;25:e41092. [doi: 10.2196/41092] [Medline: 37467015]
38. Sheehan J, Laver K, Bhopti A, et al. Methods and effectiveness of communication between hospital allied health and primary care practitioners: a systematic narrative review. J Multidiscip Healthc 2021;14:493-511. [doi: 10.2147/JMDH.S295549] [Medline: 33654406]
39. Su JJ, Yu DSF. Effects of a nurse-led eHealth cardiac rehabilitation programme on health outcomes of patients with coronary heart disease: a randomised controlled trial. Int J Nurs Stud 2021 Oct;122:104040. [doi: 10.1016/j.ijnurstu.2021.104040] [Medline: 34333211]

40. Murphy L, Moore S, Swan J, Hehir D, Ryan J. Examining the impact of video-based outpatient education on patient demand for a rheumatology CNS service. Br J Nurs 2021 Oct 14;30(18):1056-1064. [doi: 10.12968/bjon.2021.30.18.1056] [Medline: 34645352]
41. Saluky S, Bahiyah N. Revolutionizing education: unlocking the potential of asynchronous video for interactive online learning. IJEH 2023;3(2):187-196. [doi: 10.58557/(ijeh).v3i2.149]
42. Taylor ML, Thomas EE, Vitangcol K, et al. Digital health experiences reported in chronic disease management: an umbrella review of qualitative studies. J Telemed Telecare 2022 Dec;28(10):705-717. [doi: 10.1177/1357633X221119620] [Medline: 36346938]
43. Chatterjee A, Strong G, Meinert E, Milne-Ives M, Halkes M, Wyatt-Haines E. The use of video for patient information and education: a scoping review of the variability and effectiveness of interventions. Patient Educ Couns 2021 Sep;104(9):2189-2199. [doi: 10.1016/j.pec.2021.02.009] [Medline: 33741233]
44. Caetano IES, Santiago LM, Marques M. Impact of written information on control and adherence in type 2 diabetes. Rev Assoc Med Bras 2018;64(2):140-147. [doi: 10.1590/1806-9282.64.02.140]
45. Linton DN, Porteous J, Eatson H, et al. Educational booklet reinforces knowledge of osteoporosis and influences intentions to improve bone health in previously diagnosed and treated patients. Osteoporos Int 2020 Sep;31(9):1703-1711. [doi: 10.1007/s00198-020-05392-2] [Medline: 32333065]
46. Piredda M, Migliozzi A, Biagioli V, Carassiti M, De Marinis MG. Written information improves patient knowledge about implanted ports. Clin J Oncol Nurs 2016 Apr;20(2):E28-E33. [doi: 10.1188/16.CJON.E28-E33] [Medline: 26991720]
47. Laparidou D, Botan V, Law GR, et al. People with diabetes and ambulance staff perceptions of a booklet-based intervention for diabetic hypoglycaemia, “Hypos can strike twice”: a mixed methods process evaluation. BMC Emerg Med 2022 Feb 8;22(1):21. [doi: 10.1186/s12873-022-00583-y] [Medline: 35135499]
48. Reda Abdou H, Sleem WF, EL-Wkeel NS. Designing and validating educational booklet about self-directed learning for nursing students. IJN 2021;8(1). [doi: 10.15640/ijn.v8n1a3]
49. Navarro O, Escrivá M, Faubel R, Traver V. Empowering patients living with chronic conditions using video as an educational tool: scoping review. J Med Internet Res 2021 Jul 6;23(7):e26427. [doi: 10.2196/26427] [Medline: 34255671]
50. Reid KRY, Reid K, Esquivel JH, et al. Using video education to improve outcomes in heart failure. Heart Lung 2019;48(5):386-394. [doi: 10.1016/j.hrtlng.2019.05.004] [Medline: 31174893]
51. Tsabedze N, Kinsey JLH, Mpanya D, Mogashoa V, Klug E, Manga P. The prevalence of depression, stress and anxiety symptoms in patients with chronic heart failure. Int J Ment Health Syst 2021 May 12;15(1):44. [doi: 10.1186/s13033-021-00467-x] [Medline: 33980322]
52. Guidotti E, Pennucci F, Valleggi A, De Rosis S, Passino C. A longitudinal assessment of chronic care pathways in real-life: self-care and outcomes of chronic heart failure patients in Tuscany. BMC Health Serv Res 2022 Sep 10;22(1):1146. [doi: 10.1186/s12913-022-08522-0] [Medline: 36088408]
53. Owusu E, Shalaby R, Elgendy H, et al. Comparison of resilience, personal recovery, and quality of life measures pre- and post-discharge from inpatient mental health units in Alberta: analysis of control group data from a randomized trial. Healthcare (Basel) 2023 Nov 14;11(22):2958. [doi: 10.3390/healthcare11222958] [Medline: 37998451]
54. Lee J, Seo E, Choi J, Min J. Effects of patient participation in the management of daily nursing goals on function recovery and resilience in surgical patients. J Clin Nurs 2018 Jul;27(13-14):2795-2803. [doi: 10.1111/jocn.14302] [Medline: 29399909]
55. Roger VL. Epidemiology of heart failure: a contemporary perspective. Circ Res 2021 May 14;128(10):1421-1434. [doi: 10.1161/CIRCRESAHA.121.318172] [Medline: 33983838]
56. Shahim B, Kapelios CJ, Savarese G, Lund LH. Global public health burden of heart failure: an updated review. Card Fail Rev 2023;9:e11. [doi: 10.15420/cfr.2023.05] [Medline: 37547123]
57. Angkurawaranon C, Papachristou Nadal I, Mallinson PAC, et al. Scalable solution for delivery of diabetes self-management education in Thailand (DSME-T): a cluster randomised trial study protocol. BMJ Open 2020 Oct 5;10(10):e036963. [doi: 10.1136/bmjopen-2020-036963] [Medline: 33020090]
58. Volandes AE, Paasche-Orlow MK, Davis AD, Eubanks R, El-Jawahri A, Seitz R. Use of video decision aids to promote advance care planning in Hilo, Hawai’i. J Gen Intern Med 2016 Sep;31(9):1035-1040. [doi: 10.1007/s11606-016-3730-2] [Medline: 27194151]
59. Kyaw TL, Ng N, Theocharaki M, Wennberg P, Sahlen KG. Cost-effectiveness of digital tools for behavior change interventions among people with chronic diseases: systematic review. Interact J Med Res 2023 Feb 16;12:e42396. [doi: 10.2196/42396] [Medline: 36795470]
60. Adam M, McMahon SA, Prober C, Bärnighausen T. Human-centered design of video-based health education: an iterative, collaborative, community-based approach. J Med Internet Res 2019 Jan 30;21(1):e12128. [doi: 10.2196/12128] [Medline: 30698531]
61. Setiawan H, Setiawan D, Mustopa AH. Development of Android-based mobile application “Cyber Gen” for genetic counselling implementation among thalassemia patients. J Phys: Conf Ser 2021 Nov 1;2111(1):012037. [doi: 10.1088/1742-6596/2111/1/012037]

## Abbreviations

**MLHFQ:** Minnesota Living With Heart Failure Questionnaire
**QoL:** quality of life

*Edited by A Coristine; submitted 02.04.25; peer-reviewed by AM Ibrahim, X Gong; revised version received 03.08.25; accepted 05.08.25; published 05.09.25.*

*Please cite as:*
*Arofiati F, Lestari FD, Setiawan H*
*The Effect of Discharge Planning Videos and Booklets on Quality of Life Among Patients With Heart Failure: Quasi-Experimental Study*
*JMIR Cardio 2025;9:e75417*
*URL: https://cardio.jmir.org/2025/1/e75417*
*doi:10.2196/75417*

Original Paper

# MARIA (Medical Assistance and Rehabilitation Intelligent Agent) for Medication Adherence in Patients With Heart Failure: Empirical Results From a Wizard of Oz Systematic Conversational Agent Design Clinical Protocol

Nik Nailah Abdullah[1*], PhD[‡]; Jia Tang[2*], MBIS; Hemad Fetrati[3*], MCS; Nor Fadhilah Binti Kaukiah[4], BSc; Sahrin Bin Saharudin[4], MBBS; Vee Sim Yong[4], MPharm; Chia How Yen[4,5], MD

[1]School of Information Technology, Monash University Malaysia, Petaling Jaya, Malaysia

[2]Faculty of Information Technology, Monash University, Melbourne, Australia

[3]Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada

[4]Clinical Research Center (Hospital Queen Elizabeth II), Institute for Clinical Research, National Institute of Health, Kota Kinabalu, Malaysia

[5]Sabah Scientific Research Society, Kota Kinabalu, Malaysia

[‡]MARIA chatbot team

[*]these authors contributed equally

**Corresponding Author:**
Nik Nailah Abdullah, PhD
School of Information Technology
Monash University Malaysia
Jalan Lagoon Selatan, Bandar Sunway, Subang Jaya, Selangor
Petaling Jaya, 47500
Malaysia
Phone: 60 03 55146000
Fax: 60 03 55146001
Email: nik.nailah@monash.edu

## Abstract

**Background:** Nonadherence to medication is a key factor contributing to high heart failure (HF) rehospitalization rates. A conversational agent (CA) or chatbot is a technology that can enhance medication adherence by helping patients self-manage their medication routines at home.

**Objective:** This study outlines the conception of a design method for developing a CA to support patients in medication adherence, utilizing design thinking as the primary process for gathering requirements, prototyping, and testing. We apply this design method to the ongoing development of Medical Assistance and Rehabilitation Intelligent Agent (MARIA), a rule-based CA.

**Methods:** Following the design thinking process, at the ideation stage, we engaged a multidisciplinary group of stakeholders (patients and pharmacists) to elicit requirements for the early conception of MARIA. In collaboration with pharmacists, we structured MARIA's dialogue into a workflow based on Adlerian therapy, a psychoeducational theory. At the testing stage, we conducted an observational study using the Wizard of Oz (WoZ) research method to simulate the MARIA prototype with 20 patient participants. This approach validated and refined our application of Adlerian therapy in the CA's dialogue. We incorporated human-likeness and trust scoring into user satisfaction assessments after each WoZ session to evaluate MARIA's feasibility and acceptance of medication adherence. Dialogue data collected through WoZ simulations were analyzed using a coding analysis technique.

**Results:** Our design method for the CA revealed gaps in MARIA's conception, including (1) handling negative responses, (2) appropriate use of emoticons to enhance human-likeness, (3) system feedback mechanisms during turn-taking delays, and (4) defining the extent to which a CA can communicate on behalf of a health care provider regarding medication adherence.

**Conclusions:** The design thinking process provided interactive steps to involve users early in the development of a CA. Notably, the use of WoZ in an observational clinical protocol highlighted the following: (1) coding analysis offered guidelines for modeling

CA dialogue with patient safety in mind; (2) incorporating human-likeness and trust in user satisfaction assessments provided insights into attributes that foster patient trust in a CA; and (3) the application of Adlerian therapy demonstrated its effectiveness in motivating patients with HF to adhere to medication within a CA framework. In conclusion, our method is valuable for modeling and validating CA interactions with patients, assessing system reliability, user expectations, and constraints. It can guide designers in leveraging existing CA technologies, such as ChatGPT or AWS Lex, for adaptation in health care settings.

***(JMIR Cardio 2025;9:e55846)*** doi:10.2196/55846

**KEYWORDS**

heart failure; medication adherence; self-monitoring; chatbot; conversational agent; Wizard of Oz; digital health

## Introduction

### Background and Motivation

Heart failure (HF) is a global concern associated with significant morbidity and mortality [1]. Recent findings from the ASIAN‐HF registry suggest a potential shift in the HF burden from North America, Western Europe, and Eastern Europe to the Asia-Pacific region [2].

According to the ASIAN‐HF registry, within Asia, Southeast Asian patients have the highest burden of risk factors and worse outcomes than Northeast and South Asian patients [2,3]. This burden pressures individuals, their families, and the health care systems through various costs, with the most prominent being repeated hospitalizations [1]. For example, as high as 10% of hospital admissions are related to HF. The total HF costs accounted for approximately 1.8% of total health expenditure [4].

Studies show that HF's rehospitalization and mortality rates were influenced by patients' medication nonadherence [5-7]. As poor self-motivation and inadequate medication knowledge are the typical reasons for medication nonadherence, doctors and health care workers should emphasize the importance of medication adherence by constantly providing appropriate encouragement and education to patients [8,9].

Research has shown that some of these factors leading to hospitalizations are preventable by close home monitoring supported by family or nurse practitioners [6]. Nonetheless, such programs are challenging to apply in our local setting due to the limited number of specialized HF nurses who can support the wider HF patient population.

Therefore, we explore related work that uses conversational agent (CA), a type of artificial intelligence (AI) application that can be leveraged to assist in the self-monitoring of patients with HF in the following section.

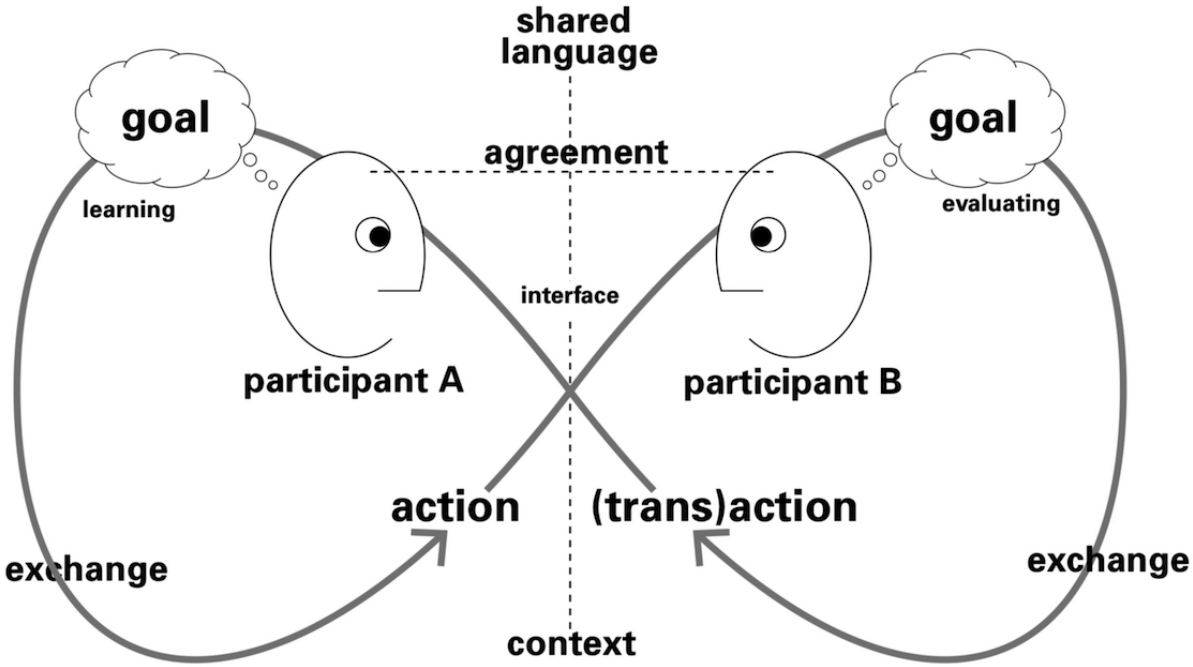
A CA is a computer program capable of understanding natural human language (in text, speech, or both forms) and responding autonomously using the same language [10]. They can be accessed through a variety of ways, such as social media platforms (eg, Facebook Messenger), websites, and smartphone apps, or deployed using stand-alone digital devices (eg, Alexa, Google Assistant, and Siri). The first CA, ELIZA, was created by Joseph Weizenbaum at the Massachusetts Institute of Technology in 1966 [11].

ELIZA was developed to converse with the users via text, imitating a psychotherapist, to fool them into believing that they were talking to a human being. Today, thanks to technological advancements in AI, CAs can handle much more complex tasks in a wide variety of fields, including finance, education, travel, and retail [12-15], and they are predicted to be used even more widely in the future [16].

Engaging in natural conversation with humans is the main characteristic of CA, and current methods refer to conversation theory (demonstrated in Figure 1 [17]), such as using advanced machine learning methods to extract users' intents from their utterances (speech) [18].

For a CA to produce natural conversations in a narrative manner, the format of the content must be outlined through rule-based workflows, templates, or intent-driven approaches to create an output. Every CA that uses a natural language system relies on narrative design, also called conversation design, to produce that output.

**Figure 1.** Simplified view of conversation theory.




Conversational design combines several disciplines, including copywriting, user experience design, interaction design, visual design, motion design, and, if relevant, voice and audio design. Conversation design not only requires using natural conversational language but also creates logically sound conversational flow and design specifications that capture the entire user experience. More recently, machine learning capabilities have been used in CA to provide the ability to learn from the data so that an adaptable context of responses can be provided to the users.

There are several ways to generate the responses. First, is the rule-based method in which the CA produces a response by selecting it from a pool of predetermined responses either following simple rules to match phrases or identifying specific keywords in the text [19].

The second type is the generative-based CAs, which use AI algorithms to develop a contextual response informed by the system's previous and ongoing learning [20].

Rule-based CAs allow developers greater control over the conversation content and flow, which is a useful feature when developing CAs for health care. By contrast, AI algorithms, particularly neural networks, may develop decisions that are not explainable or understood by the end user, referred to as the *black box* [20]. In health care settings, the *black box* effect may lead to biased or erroneous decision-making and patient harm which is highly dependent on the type of algorithms used to learn and generate the responses.

Therefore, in our work, we choose to develop a rule-based CA, given that it will allow developers better control and transparency in the responses.

Researchers have effectively innovated the application of CA in the digital health (DH) area, covering functions such as scheduling doctor appointments, monitoring medication intake, checking symptoms, diagnosing, providing treatment plans, and helping patients with rehabilitation [21-24]. DH has a broad scope that includes categories such as mobile health, health information technology, wearable devices, telehealth and telemedicine, and personalized medicine [7].

There are existing applications developed for supporting patients with HF. CARDIAC is a human-centered conversational assistant that helps patients with HF monitor their health status through reminders, question answering, relevant data collection, and generating data tendencies and personal health records [25]. Another CA, DIL, improves the self-care and quality of life of patients with HF by motivating them to adhere to a healthy lifestyle, including a controlled diet, a continuous medication routine, and regular exercise [26]. As a medication advisor, CARMIE speaks in Portuguese and interacts with patients with HF in real time to provide quality answers to medication-related questions according to its knowledge representation model and patients' prescriptions [27].

Based on our literature review [10-12,26,28-31], the existing CAs in the HF area concentrated on developing functional features' effectiveness and accuracy. However, no study has specifically displayed a method for building agents' natural language–based conversations to encourage and educate patients with HF about medication adherence, nor a standard for evaluating this type of CA design early in the development stage as a DH solution.

Therefore, our study aims to adopt established design methods and conceive them into a systematic method that uses a clinical

observational study protocol. We use observational study protocol to produce new knowledge in improving conversational design, examine acceptability, and reduce uncertainties in the harmful effects of using CA in medication adherence. It will fill the gap of the existing studies in the DH domain in designing a CA (or chatbot) that encourages and educates patients about medication adherence.

## Prior Work

### *Overview*

In the following subsections, we will review the prior work in related research studies.

### *Designing a CA Agent With Human-Likeness Attributes*

To fill the gap in the existing studies and strategically motivate patients to change medication adherence behavior, we searched for suitable psychological theories to support our CA dialogue. Adlerian psychoeducational therapy emphasizes that encouragement is the key to achieving an individual's growth and development [13]. Developed by Alfred Adler [32], the approach states that the motivation of an individual's behavior change can be goal oriented and related to one's relationship with others and contributions to society [14]. This therapy aims to help individuals identify their mistaken beliefs in their capabilities and apply appropriate improvements to reinforce their strengths and compensate for their weaknesses. It encourages individuals to regain their confidence in achieving their goals. The therapy is widely used in mental health treatment for anxiety, depression, behavior disorders, mental disorders, and career encouragement [15]. Adlerian psychologists encourage their patients by using therapeutic skills. For instance, they enhance patients' self-efficacy and affirm patients' capabilities and potentials by narrating other patients' successful experiences to build good examples. They help patients recognize and believe in their strengths, resources, progress, and positive sides of life experiences and encourage them to keep striving toward their goals [16].

The storytelling method to encourage individuals to learn how relevant peers have successfully solved a similar problem is also conceptualized in Social Cognitive Theory [33,34]. Being expanded by Albert Bandura [35], Social Cognitive Theory studies individuals' behavior change through the impact of individuals' experiences, the achievements of others, and the influences from surroundings [36]. The theory believes that an individual could learn similar behaviors from observing the successful experiences of others [37].

The Tripartite Encouragement Model is a psychological framework that combines the insights of encouragement, verbal persuasion, and character strength and virtues [16]. The Tripartite Encouragement Model introduces the concept of effective encouragement to optimize the positive influences of encouragement to recipients. An encouragement message could effectively motivate recipients' self-efficacy by emphasizing their progress rather than pointing out their distance apart from the target. Highlighting the process-oriented factors is another way to improve the effectiveness of encouragement, such as emphasizing the recipient's positive effort, attitude, and feelings. Cialdini and Sagarin's [18] principles of interpersonal influence contain psychological persuasion strategies to trigger individuals' acceptance of requests while hesitating. The principle of commitment and consistency states that individuals tend to accept a request consistent with their committed position [18]. The 4-wall technique asks individuals several easy-to-say-"yes" questions first, then leads them to comply with the final crucial request [38]. The principle of reciprocity demonstrates that individuals tend to accept a request if requestors offer a concession [18]. The reciprocal concession procedure significantly reduces the requested content after the initial request gets rejected, which could make the new request more acceptable [39].

Anthropomorphism, or human-likeness, is a phenomenon that also occurs in human-technology interaction contexts. It is used to enhance user experience in chatbots. This approach is typically implemented through the CA or chatbot's visual representation, such as an illustration, image, or animated avatar, alongside a persona that defines various humanlike characteristics, including sex, gender, education, race, and age [40,41]. These features are often selected to reflect the target audience, such as an avatar having a similar skin tone, wearing local attire, or having a common local name [42]. Additionally, conversation style plays a crucial role, with the use of slang, local accents, and culturally appropriate vocabulary tailored to the users' demographic [40]. Another significant factor in shaping a chatbot's humanlike persona is its social role. For example, adopting a peer persona or an expert persona (eg, a doctor) has been shown to be effective, particularly in medical-related chatbots [40].

The existing design guidelines for CAs explain that similarity attraction significantly impacts users' acceptance of the system because individuals tend to apply human-human interaction to engage with virtual agents [43]. Individuals prefer to engage with those with similar experiences or interests, and the similarities could create more conversations to establish relationships and trust [44]. Existing studies also suggest that the human-likeness of the CA is essential [43]. Human beings spontaneously mix emotions and languages to display their feelings and reactions during face-to-face conversations. Emojis can display speakers' emotions and optimize the chatting experiences during text-based online communication [45]. Some studies recommend adding an intentional pause between messages sent and received to generate a natural feeling as chatting with a human [46]. The pause will also allow users to think and type their responses [47]. When applying encouragement and education strategies, the credibility appeal could be enhanced by providing reliable evidence of the information to users [48]. Furthermore, people tend to trust an individual with a consistent personality that indicates one's capability, predictability, and reliability [43]. The patterns in language use could reveal one's personality [49]. Moreover, finding the right balance of anthropomorphism—without overdoing it, which can diminish the sense of human-likeness—has been shown to increase user engagement, compliance, satisfaction, and the intention to reuse chatbots [50].

In applying an agent-based concept in modeling CA, protocols play a central role in agent communication with humans or another CA. A protocol specifies the rules of interaction between 2 or more communicating agents by restricting the range of allowed follow-up utterances for each agent at any stage during a communicative interaction (dialogue). Such a protocol may be imposed by the designer of a particular system or it may have been agreed upon by the agents taking part in a particular communicative interaction before that interaction takes place [51].

### *Wizard of Oz Procedure in the Elicitation of Requirements and User Experience*

Wizard of Oz (WoZ) is a well-established method for simulating the functionality and user experience of future systems, where humans simulate all or part of the behaviors and functionalities of an automated system [52,53]. Using a human wizard to mimic certain operations of a potential system is particularly useful in situations where extensive engineering effort would otherwise be needed to explore the design possibilities offered by such operations [53].

The term "Wizard of Oz (WoZ)" was first coined by John Kelley [54], who used this technique to simulate a calendar application that could be operated via natural language input [53]. The method was also occasionally referred to as "Pay No Attention to the Man Behind the Curtain" and "OZ paradigm" [53,55]. Over time, the use of WoZ expanded beyond the use of simulating text-based interfaces to include interfaces involving speech, gesture, facial recognition, and multimodal user interactions [53,56-58].

There are several key uses of the WoZ method for designing interactive systems. One major application is in interaction design, where WoZ is used to explore human-computer dialogues and interaction strategies. Additionally, WoZ is used to collect text and speech corpora (ie, eliciting requirements), which aids both interaction design and engineering work by training and fine-tuning technology components. A third key use involves employing WoZ to develop early prototype technology components, allowing for the evaluation of system performance in specific application areas without the need for full-scale engineering efforts. Overall, these uses fall into 4 broad categories: exploring interaction strategies, designing dialogues, collecting corpora, and evaluating system components [53].

In recent years, researchers have utilized WoZ for various purposes within these categories, such as building a data set to create a virtual assistant for helping programmers use application programming interfaces [59], simulating autonomous driving cars [60,61], developing drive-assist features [62], conducting virtual reality elicitation studies [63], and creating a mixed reality game [64].

In our study, we use the WoZ method for 2 main objectives. First, to simulate the Medical Assistance and Rehabilitation Intelligent Agent (MARIA) prototype to validate and improve our use of Alderian theory in designing the CA's workflow for medication adherence. Second, to test and improve the overall user experiences using MARIA, which engages users in adherence to medication.

### Goal of Study

The goal of our study is to conceive a design method for developing CA for patients' use in medication adherence, using design thinking as the main process for gathering requirements, prototyping, and testing.

We apply our design method in the ongoing development of MARIA, a rule-based CA.

The end goal of the study is to identify improvements in the functionality and dialogue construction of MARIA. This could be applied to leverage existing technologies that use CA or chatbot, such as ChatGPT or AWS Lex, to adapt it within a health care setting.

In this paper, we report on the results of our observation study protocol applying our design method for CA development.

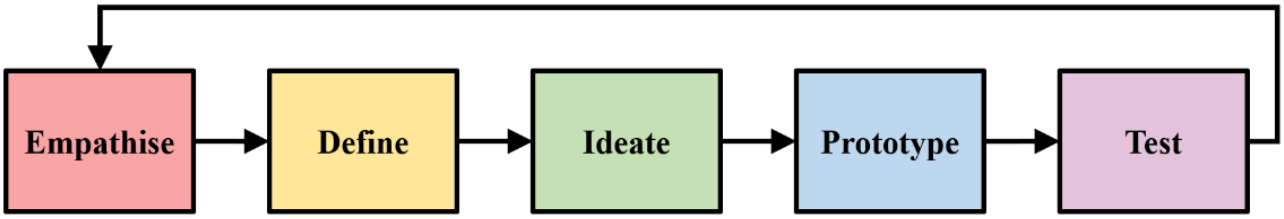
## Methods

### Design Thinking Processes

#### *Methodology Processes*

The design thinking methodology consists of 5 processes (phases) [65]: empathize, define, ideate, prototype, and test, as shown in Figure 2.

**Figure 2.** Design thinking methodology.




The process can be nonlinear and iterated until the best solution to the problem is achieved [66]. In our research, we conducted 1 iteration of the design thinking process to improve our prototype design.

#### *Empathize*

Constructing empathy to understand the stakeholders and their problems is essential in human-centered consideration and is the core of the design thinking process [67]. In our research, we conducted the steps outlined in Textbox 1 to gather detailed information to understand the problem and stakeholders' needs better.

**Textbox 1.** Steps to gather information to understand the problem and stakeholders' needs better.

- Review of the current state of the system

We reviewed the previous achievements of Medical Assistance and Rehabilitation Intelligent Agent's (MARIA) design to observe the relevant context, including the tasks accomplished by the Monash research team in this project [33,68].

- Work practice observations and interviews

As MARIA aims to perform as a personal nurse assistant to motivate patients about medication adherence, we studied the work procedures for managing patients with heart failure (HF) in Malaysian cardiac centers. We use ethnographic studies and interviews as a method to gain insights into the work practices in the management of patients with HF [33].

- Design thinking meeting

We organized a design thinking meeting to collect stakeholders' requirements and practice knowledge about encouraging and educating patients with HF to adhere to their medication. We refer to the requirements method in the work by Abdullah et al [33] where several iterations of meetings take place.

The meeting involves direct and indirect stakeholders, those who will be using it directly (patients) and those who are part of the patient management team (pharmacists and specialists). Specifically for our work, we involved the supervisor from Monash Malaysia as the project lead, at least 3 medical doctors from the Malaysian cardiac centers, 2 pharmacists, 3 developers, and 1 student researcher from Monash Australia as the MARIA conversational agent designer. The meetings were conducted iteratively until all team members reached common ground on the pain points of HF management, as well as the challenges faced by health care practitioners in ensuring medication compliance in these patients. Every meeting was recorded for further analysis by the researcher and validated by the team.

### *Define*

Based on the requirements of stakeholders' needs and the research context, the "Define" stage identifies the problem and the factors contributing to this problem [67]. We applied the thematic, qualitative analysis approach to capture stakeholders' essential requirements and the core issue [69]. We created the edited transcription to omit the unnecessary content in the recorded meeting conversations to help us retain the recording quality and capture the critical information in the collected data [70]. We marked the latent codes in our meeting transcription to demonstrate the underlying themes from the interpretative level [69]. Then, we analyzed and categorized the thematic codes to define the critical problem and stakeholders' expectations in MARIA's expanding design.

### *Ideate*

The conceptual solution to the defined problem is generated in the ideate phase, and the brainstormed outcomes are the potential source for building the prototype [66]. We integrated the literature review of the relevant studies, the context learning of the cardiac center's work procedures, and the thematic analysis of stakeholder's requirements, and then visually demonstrated our design concept in the MARIA Interaction Protocol for Motivating Patients. We used a workflow diagram to display our protocol. The diagram can illustrate the step-by-step procedure for completing a task in a logical sequence, define how information and responsibility are transferred between parties during the task, clearly indicate the beginning and end of the process, and display parallel paths reflecting the consequences of different decisions or alternative options [71]. Our protocol contained the set of activities that MARIA should carry out and follow during the interaction with patients with HF. The activities were designed to ensure MARIA performs the role of personal nurse assistant to encourage and educate patients about medication adherence from home and reduce rehospitalizations and medical staff's workload.

### *Prototype*

A prototype is a quick and cost-saving conceptual model built to obtain valuable user feedback for further optimization considering the final product's practical application [67]. It leads the design closer to the final solution [66]. Based on our proposed protocol, we prototyped the conversational templates using the decision tree method. This method is commonly adopted in designing the data-mining algorithm for predicting multiple target variables [72]. We designed our decision-tree templates to suit the future programming of the MARIA conversational system [68]. Encouragement and education strategies were included in the conversational templates to enhance patients' confidence in medication adherence. The design also covered the reinforcement of MARIA's human-likeness and reliability to enhance patients' user experience and trust for the long-term use of the MARIA application.

### *Test*

The test stage provides another opportunity to apply empathy by comparing the user feedback and the initial understanding of the requirements. It evaluates whether the defined problem has been successfully addressed and delivers the information for refining the prototype [66].

We use an observational study protocol to design the WoZ method and a user satisfaction scoring test at this step.

WoZ was used to simulate MARIA to validate and improve our use of dialogue designs. The user satisfaction scoring test, by contrast, was used to evaluate the engagement of patients with the MARIA prototype (Multimedia Appendix 1).

## The WoZ Method for the Observational Study Protocol

Our conceived WoZ in an observational study protocol was designed to simulate the interaction of MARIA with participants, aiming to validate (testing) and refine template responses (ie,

CA's workflow dialogue) while gathering user experience feedback.

Given that the aim of using WoZ was ultimately to improve the design of a rule-based CA, we did not control for participants' beliefs about whether they were interacting with a real person or whether the study procedure (ie, the MARIA prototype) was successful. Instead, participants interacting with MARIA believed it was autonomous. Our researcher (CHY), acting as the wizard, operated MARIA from another room.

The number of participants varies from one work to the other with no consensus on the ideal number of participants when used in a WoZ method. For example, the work of Bonial et al [73] involved 10 participants in the study. On the other hand, Nielsen and Norman's [74] recommendation for usability testing, which the WoZ also falls into, required 5 participants to test. By contrast, in requirements elicitation [75], there are no specific guidelines for the number of persons required; it can vary from 2 to 12 persons.

Given that there is no agreement on the number of sample sizes, we follow a qualitative study recommendation of 20 samples [76] as an initial sample size. Furthermore, because the protocol is designed as an interactive process, researchers may stop to recruit further sample size when analysis suggests that data are saturated (ie, not many differences in the responses at a certain point).

## Ethical Considerations

MARIA_PRO_VER_3_190122 is registered with the Malaysia Medical Ethics Committee. The Medical Research Ethics Committee, the Ministry of Health Malaysia, approved the study with the registration number NMRR-21-1388-60672 (IIR). Patients provided informed consent before their involvement in the study and consented to use their data for analysis. The patients were provided compensation after completing the WoZ study.

## Privacy and Confidentiality Protection

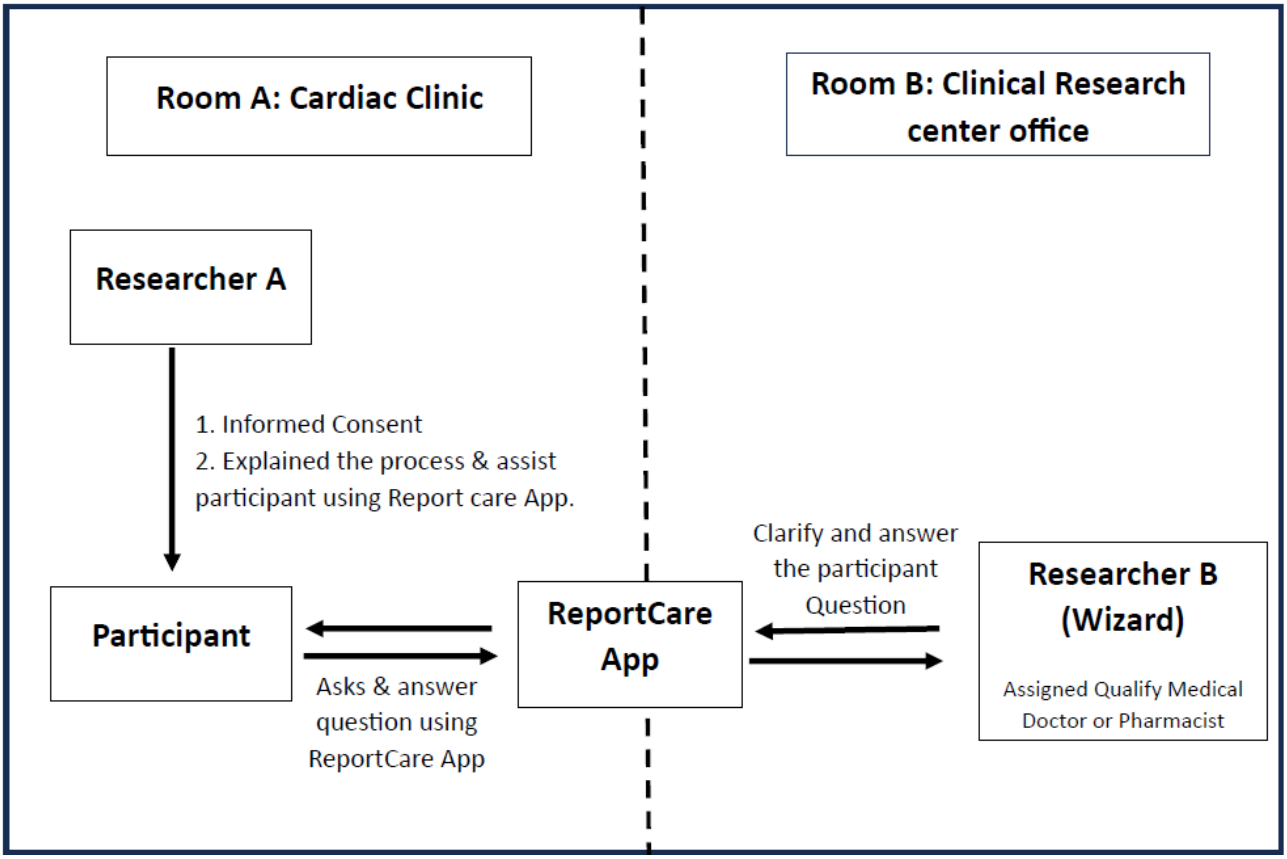
Participant names for this research have been deidentified and linked only with a study identification number. Therefore, the research did not identify the participant's identity and instead used anonymized identification numbers on all the data sets. All data are stored in Monash University Malaysia REDCap secured cloud and kept for 3 years. Participants can write to the investigators to request access to study findings.

## Study Procedure

During the recruitment and study period, there were 2 researchers, researcher A and researcher B, each located in separate facilities. Researcher A was based in the cardiac clinic, whereas researcher B operated from the Clinical Research Center (CRC) office. Participants were assigned to the cardiac clinic with researcher A. Researcher B worked from the CRC office (refer to Figure 3).

The study protocol allowed only 1 participant at a time in each room, with each session being conducted sequentially, 1 participant following another. Textbox 2 provides an explanation of the roles and responsibilities of the researcher and participant. Part A details the roles of researcher A and the participant, while part B outlines the responsibilities of researcher B.

**Figure 3.** Overall Wizard of Oz study procedure.



**Textbox 2.** Roles and responsibilities of the researcher and participant.

- Part A: role of researcher A and participant/setting: room A—cardiac clinic
- Informed consent process:
  - Researcher A explained the details of the research and the participant signed the consent.
  - The participant is provided with a unique ID for deidentification purposes.
- Explanation of the process and assisting participants in using the web app:
  - The participant will be seated in a room and given a smartphone with the web app preinstalled.
  - Researcher A will explain how to use the web app and Medical Assistance and Rehabilitation Intelligent Agent (MARIA), the messaging chatbot as a self-management tool, in a home setting.
  - The participant will log into the web app using the unique ID provided.
- Given scenarios:
  - Researcher A gives a set of written scenarios to participants (for participants to recall their usual symptoms or signs that they experienced) and the common questions or clarification participants would like to ask MARIA related to the given scenario.
  - The participants will respond with their questions based on the scenario using the web app messaging feature.
- Part B: role of researcher B (to role-play the wizard) delegated to a qualified medical doctor and pharmacist/setting: room B—Clinical Research Center office
- Researcher B will be provided with the participant ID and basic information (sociodemographic and medication history).
- Researcher B will refer to the Heart Failure Clinical Practice Guidelines [23] and the Pharmacy Practice and Development Division, the Ministry of Health Malaysia [77], and the Protocol for the Medication Therapy Adherence Clinic [24]. In particular, the researcher will follow:
  - The workflow on therapy medication protocol adherence for furosemide titration, including management of side effects.
  - The workflow for general inquiries on the medication side effects of furosemide and beta-blockers [78].
  - The workflow on the management of symptoms and signs.
- According to the standard workflow, researcher B will respond to participants via the messaging chatbot provided in the ReportCare app.
  - Pharmacists and medical doctors will respond to drug- or clinical-related questions such as medication titration, drug dosage, frequency, side effects, and drug interaction.

## Recruitment

Study participants were recruited from the Hospital Queen Elizabeth II, Sabah in Malaysia. The participant recruitment process was from June 2022 to November 2022.

The recruitment process followed the Malaysian Good Clinical Practice guidelines. The participants for this study were identified by CHY (principal investigator) at the HF clinic. During the consultation, the investigator explained the study to the patients and provided the consent form. If the patient fulfilled the inclusion and exclusion criteria, they were given sufficient time to read, discuss the study, and ask any questions. All questions were answered by the investigator. After addressing the patient's concerns, the patient signed the consent form.

## Study Population

The study population included patients with chronic HF who were currently being followed up at the Cardiology Department Outpatient Clinic in Hospital Queen Elizabeth II. The inclusion criteria were: (1) age above 18 years, (2) diagnosis of chronic HF for at least one year, (3) history of symptomatic HF, (4) ability to write and speak Malay and English, (5) ability to type and use mobile app messaging, and (6) ability to comply with the protocol.

The exclusion criteria were as follows: (1) the presence of a clinical condition that would interfere with participation in the interview and (2) mental or legal incapacitation preventing the patient from providing informed consent.

## Sample Size

Typically, the sample size is small at the beginning, as the goal is to explore the system. With each improvement, the process continues until an acceptable usability score or set of requirements is achieved [73-75].

As stated in the "The WoZ Method for Observational Study Protocol" section, given the lack of agreement on sample size, we follow a qualitative study recommendation of 20 samples [76].

We use usability scoring as a quantitative standard to determine the acceptability of the system's design before proceeding with implementation. Hence, for the initial sample size, we used a convenience sampling method, recruiting a minimum of 20 patients for the study.

- Ten participants can speak and write the Malay language.
- Ten participants can speak and write the English language.

## Study Duration

The total time required for each participant to participate in the study was a maximum of 1 hour.

## Wizard Protocol

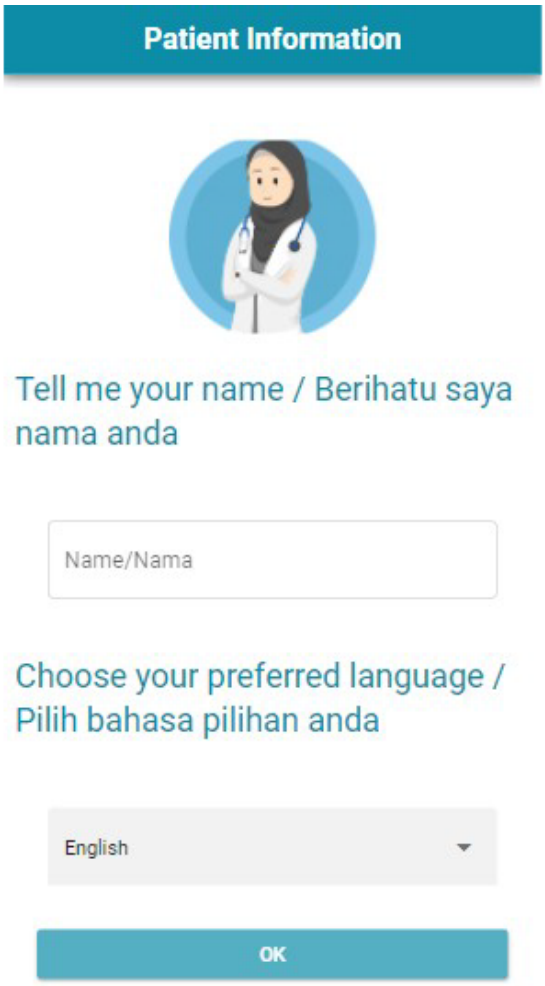
### *Overview*

Below, we share an excerpt from MARIA's workflow protocol for goal setting, daily monitoring, and goal completion.

### *Wizard Preparation*

The wizard (researcher B) launched the web app (Figure 4) before the patient, entered "MARIA" as the name, and selected either English or Malay based on the patient's preferred language. The participant then waited to launch the web app (refer to participant protocol). The wizard entered the participant's name, after which the web app redirected to the chatbox, where the participant entered their name(s).

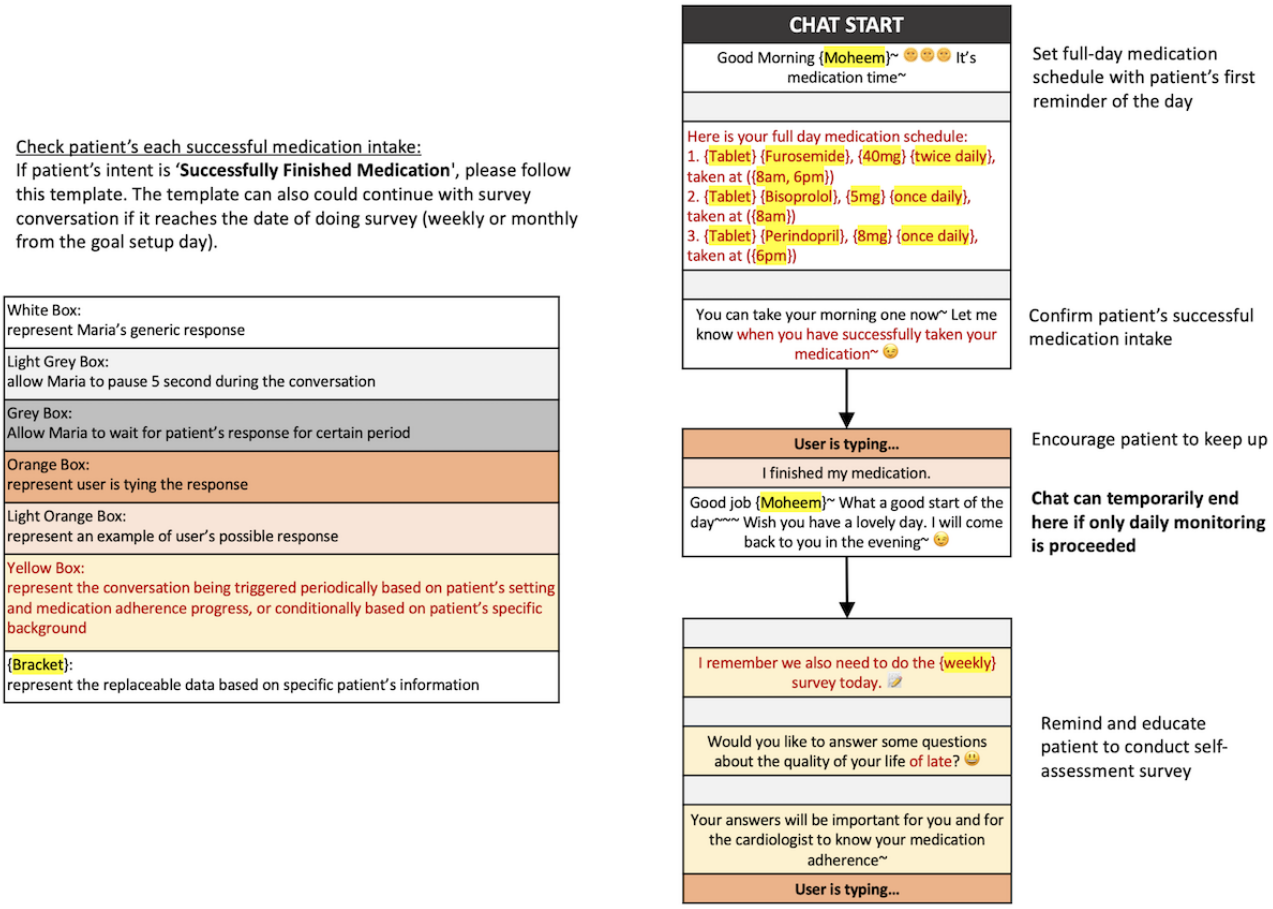
**Figure 4.** Screenshot of the app displaying the log-in interface including the language selection feature.




### *Conversation Protocol*

In this study, the participant will ask questions based on the conversation flowchart (Figure 5). If the question follows the predefined flow, researcher B (wizard) will respond or ask a follow-up question accordingly. However, if the question or response deviates from the flow, researcher B (wizard) will intervene, providing an appropriate response or asking a relevant question to steer the conversation back on track. This intervention ensures that the discussion remains focused and addresses any inquiries outside the predefined flow. Researcher B (wizard) will continue following the conversation flowchart and await the participant's responses.

**Figure 5.** Overview of the conversation protocol as followed by the Wizard throughout the study.




### *Participant Protocol*

Researcher A is responsible for obtaining participants' consent and collecting their basic demographic and medical history information, which is then provided to the wizard (researcher B) for further analysis.

Researcher A also assists participants in launching the web app on their mobile devices. Once participants enter the chat room, they can ask questions or respond using the web app interface.

Before participants begin their conversation with the wizard, researcher A explains the research process, which is divided into 3 parts: part 1 (goal setting), part 2 (daily monitoring), and part 3 (goal completion). Each part is explained in detail to the participant.

In part 1 (goal setting), researcher A highlights the importance of goal setting, while the wizard (researcher B) follows the predefined flowchart to assist participants in setting up medication reminders and emergency contacts.

In part 2, researcher A presents scenarios related to medication adherence, such as remembering or forgetting to take medication. Participants respond to these scenarios, and the wizard (researcher B) provides appropriate replies based on their answers.

In part 3, the wizard (researcher B) follows the conversation flowchart to ask participants about their quality of life and updates the relevant information accordingly.

## Conversation Analysis

We developed a coding guideline for analyzing the utterances, as detailed in Textbox 3.

The researcher tested the coding guideline before providing it to the clinical researcher, who then used it to analyze the collected data from the study participants.

**Textbox 3.** Coding guideline.

Objectives of coding

- To identify speech act verbs of each utterance
- To identify turn-taking
- To identify which workflow was used to map each utterance
- To annotate the workflow part that has been modified

Instructions

- Follow the sample provided for annotating each individual's chat logs.

Workflow

Each utterance is mapped to the workflow that was used by the wizard as follows:

- If it is not in the workflow, simply annotate with N/A (not applicable)
- If it is part of the workflow, simply annotate the corresponding workflow reference (eg, "Workflow: Daily Monitoring")
- If it is part of the workflow but was modified during the study, add the remark "Modified" in the remark column.

Speech act definition and example of annotation

A speech act is an utterance that serves a communicative function. We perform speech acts when we offer an apology, greeting, request, complaint, invitation, compliment, or refusal. A speech act may consist of a single word, such as "Sorry!" to express an apology, or multiple sentences, such as "I'm sorry I forgot your birthday. It just slipped my mind." Speech acts occur in real-life interactions and require not only linguistic knowledge but also an understanding of appropriate language use within a given cultural context.

Here are some examples of speech acts we use or hear every day:

*Greeting:* "Hi, Eric. How are things going?"

*Request:* "Could you pass me the mashed potatoes, please?"

*Complaint:* "I've already been waiting three weeks for the computer, and I was told it would be delivered within a week."

For the speech act definition, we refer to the work of Vanderveken [79].

Topic

The topic, in essence, is what is being communicated in a sentence. You may use the topics identified by the template. If none of the provided topics fit the chat you are analyzing, you may define a new topic.

Turn-taking definition and analysis

- Turn-taking occurs in a conversation when one person listens while the other speaks. As the conversation progresses, the roles of listener and speaker are exchanged back and forth in a cyclical manner.
- Analyzing turn-taking is essential to assess whether both participants are engaged in communication. It can be examined using different units of measurement, such as adjacency pairs, continuing turns, and intervention turns.
- For our dialogue modeling, we use adjacency pair turn-taking as the unit of analysis. Adjacency pairs consist of 2 utterances produced by different speakers. To form an adjacency pair, there must be at least two speakers. In adjacency pairs, the first utterance—known as the first pair part—requires a response, while the second utterance—known as the second pair part—serves as the response to the first.

Here are some examples:

**Question and answer**

*Speaker 1:* "Where's the milk I bought this morning?"

*Speaker 2:* "On the counter invitation."

**Invitation and Acceptance**

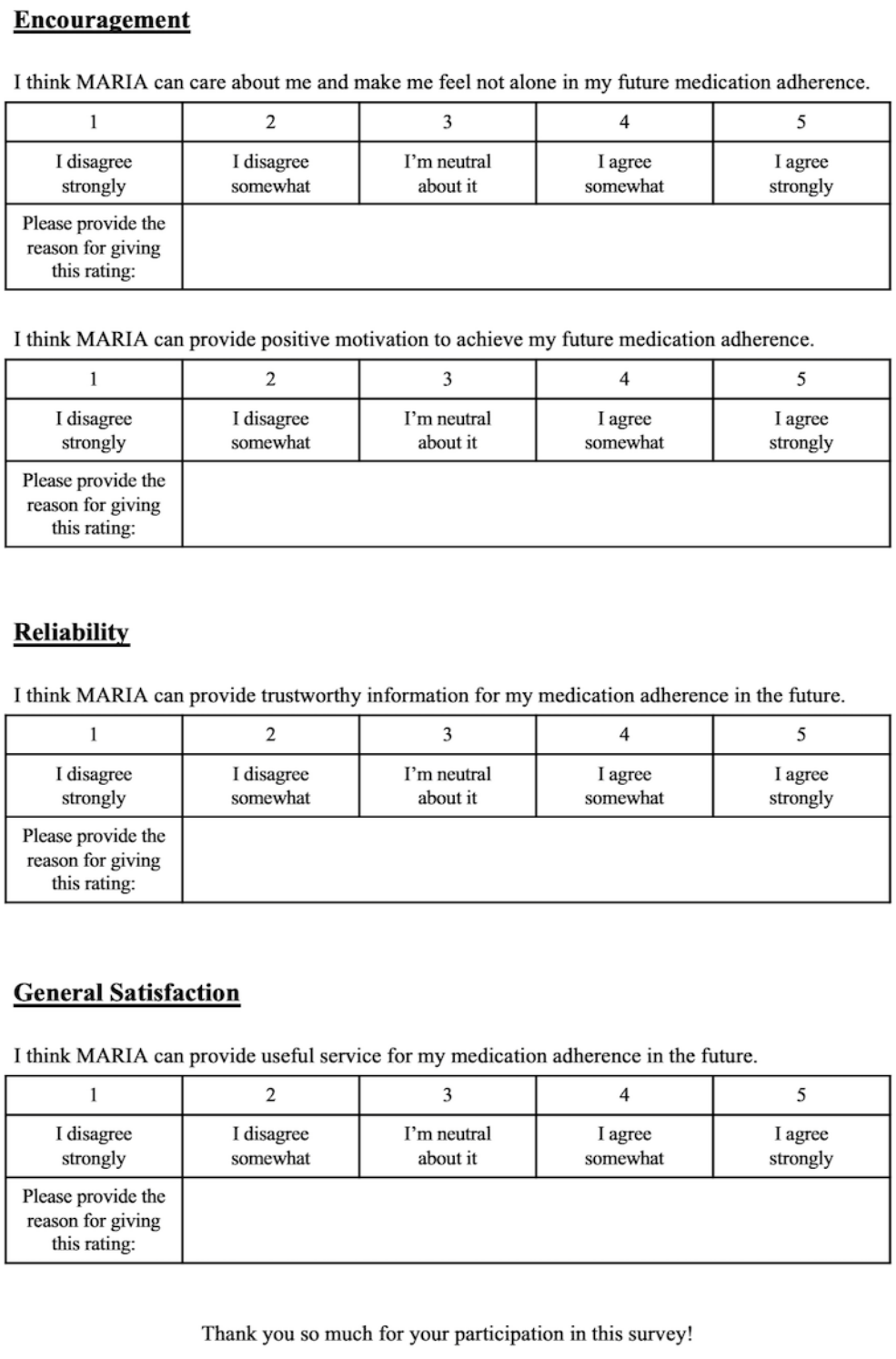
*Speaker 1:* "I'm having some people to dinner on Saturday, and I'd really like you to come."

*Speaker 2:* "Sure!"

## User Satisfaction Scoring Test

We used Hoffman et al's [80] evaluation of user trust in AI systems. Our questionnaire includes Likert-scale questions rated from 1 to 5, where 1 represents "I disagree strongly" and 5 represents "I agree strongly." Additionally, we included open-ended questions to understand the reasons behind the given ratings. The questionnaire focuses on evaluating our conversational template design from various aspects (Figure 6), including human-likeness.

**Figure 6.** An excerpt from the usability evaluation survey.



**Encouragement**

I think MARIA can care about me and make me feel not alone in my future medication adherence.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| I disagree strongly | I disagree somewhat | I'm neutral about it | I agree somewhat | I agree strongly |
| Please provide the reason for giving this rating: | | | | |

I think MARIA can provide positive motivation to achieve my future medication adherence.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| I disagree strongly | I disagree somewhat | I'm neutral about it | I agree somewhat | I agree strongly |
| Please provide the reason for giving this rating: | | | | |

**Reliability**

I think MARIA can provide trustworthy information for my medication adherence in the future.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| I disagree strongly | I disagree somewhat | I'm neutral about it | I agree somewhat | I agree strongly |
| Please provide the reason for giving this rating: | | | | |

**General Satisfaction**

I think MARIA can provide useful service for my medication adherence in the future.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| I disagree strongly | I disagree somewhat | I'm neutral about it | I agree somewhat | I agree strongly |
| Please provide the reason for giving this rating: | | | | |

Thank you so much for your participation in this survey!

For human-likeness, which encompasses MARIA's natural human language use, personality consistency, and expressed emotions, we define the criteria used for usability scoring.

- Educational strategies: Evaluate MARIA's effectiveness in tutoring patients on completing daily medication intake and providing appropriate knowledge to clarify medication use and side effects.
- Encouraging strategies: Assess MARIA's ability to offer care, support, and positive reinforcement to motivate patients toward medication adherence.
- Reliability: Reflects patients' trust in the accuracy of the information provided by MARIA during interactions.
- General satisfaction: Captures the overall impression of MARIA's conversations and their applicability.

# Results

## Evaluation of MARIA's Conversational Design and Its Implications for Medication Adherence

The evaluation outcomes indicate that our conversational template design generally met the needs of stakeholders, including end users, patients, and pharmacists. MARIA's natural language interactions, along with its encouragement and education strategies, are expected to support medication adherence among patients with HF in the future. However, the study also highlighted concerns regarding system liability and raised discussions on the extent to which MARIA should provide educational content on medication interactions and side effects in response to patient inquiries.

## Evaluation

### *Coding Analysis*

Each logged utterance was transferred into an Excel sheet (Microsoft Corporation). Independent coders (ie, clinical researchers) conducted the coding analysis based on the provided instructions (Multimedia Appendix 2). An example of the coding analysis is presented in Multimedia Appendix 3.

On average, study participants engaged in 30 interactions with the wizard, with a turn-taking ratio of 4:1 between the wizard and participants per topic. This pattern indicates that participants primarily engaged in question-answer exchanges with the wizard. The topics and speech acts used in the dialogue aligned with psychoeducational therapy theory, as evidenced by annotations of speech acts such as suggestions, support, and applause. However, having the wizard simulate MARIA revealed gaps in the workflow, including challenges in addressing negative responses, the appropriate use of emoticons, and the system's feedback mechanism during turn-taking delays.

Regarding topics, patients were most interested in asking about medication interactions and side effects. However, given MARIA's high average turn-taking per study participant, patients provided feedback suggesting that chat messages should be more concise—ideally limited to a single sentence. Longer messages often cause patients to lose track of the topic, requiring them to re-read the content for clarity.

### *Usability Scoring*

Table 1 presents the evaluation results for the usability scoring of the MARIA CA design, including demographic data of the study participants.

The human-likeness of interactions with MARIA received a median score of 4.75 out of 5. However, MARIA's personality scored lower, with a median of 3.8. In terms of natural language use, patients generally felt that conversing with MARIA resembled real human communication (question 1). One participant noted, "I am aware that I'm chatting with an AI. However, most responses were similar to what I would expect from a human."

However, MARIA's demonstration of personality and emotions (question 2) received the lowest rating in the evaluation. While the designed conversations made patients feel friendly and cared for, one patient noted a lack of distinct character in MARIA as a health assistant.

Regarding guiding patients to follow the medication routine (question 3), all fictional patients believed that MARIA's tutoring strategy would effectively support future medication adherence.

Feedback indicated that the educational content provided by MARIA was clear and easy to understand, with its knowledge-sharing approach helping patients learn about medication functions (question 4).

Additionally, in terms of encouragement strategies, fictional patients confirmed that MARIA's conversations were highly encouraging, fostering a sense of support and assisting with medication adherence (question 5).

"It is a good feeling if you open your phone, and someone (AI) keeps reminding you about your medication," one patient commented, highlighting MARIA's role in fostering adherence. Additional feedback reinforced MARIA's supportive nature, with remarks such as "MARIA is supportive of me, and I feel motivated every day" and "MARIA is very perseverant" (question 6).

Regarding reliability (question 7), 1 patient expressed trust in MARIA for medication management, while another noted the need to confirm information with a doctor. Despite this, MARIA received an average satisfaction score of 4.5 (question 8), with patients affirming its effectiveness in reminding them to take their medication on time.

From a patient safety perspective, the wizard, played by the pharmacist, played a crucial role in defining the extent to which a CA could communicate on behalf of a health care provider regarding medication adherence. Initially, the study included a workflow for educating patients about medication side effects. However, concerns arose about the implications of automating responses by retrieving drug side effect information from web-based sources. Based on these concerns, the decision was made to remove the workflow for medication side effects to ensure accuracy and patient safety.

**Table 1.** Participants' demographic data and usability evaluation results.

| Demographic | Values |
|---|---|
| **Sex, n** | |
| Male | 15 |
| Female | 5 |
| Age (years), mean | 49 |
| **Human-likeness** | |
| **I think MARIA[a] can talk like a real person, n** | |
| I disagree strongly | 0 |
| I disagree somewhat | 0 |
| I am neutral about it | 0 |
| I agree somewhat | 5 |
| I agree strongly | 15 |
| **I think MARIA can show her personality and emotion during the conversation, n** | |
| I disagree strongly | 2 |
| I disagree somewhat | 2 |
| I am neutral about it | 2 |
| I agree somewhat | 6 |
| I agree strongly | 8 |
| **Education** | |
| **I think MARIA can guide me to complete my daily medications in the future, n** | |
| I disagree strongly | 0 |
| I disagree somewhat | 0 |
| I am neutral about it | 1 |
| I agree somewhat | 4 |
| I agree strongly | 15 |
| **I think MARIA can remove my misunderstanding about medication use and side effects, n** | |
| I disagree strongly | 0 |
| I disagree somewhat | 0 |
| I am neutral about it | 4 |
| I agree somewhat | 7 |
| I agree strongly | 9 |
| **Encouragement** | |
| **I think MARIA can care about me and make me feel not alone in my future medication adherence, n** | |
| I disagree strongly | 0 |
| I disagree somewhat | 0 |
| I am neutral about it | 2 |
| I agree somewhat | 8 |
| I agree strongly | 10 |
| **I think MARIA can provide positive motivation to achieve my future medication adherence, n** | |

| Demographic | Values |
|---|---|
| I disagree strongly | 0 |
| I disagree somewhat | 0 |
| I am neutral about it | 2 |
| I agree somewhat | 8 |
| I agree strongly | 10 |
| **Reliability** | |
| **I think MARIA can provide trustworthy information for my medication adherence in the future, n** | |
| I disagree strongly | 0 |
| I disagree somewhat | 0 |
| I am neutral about it | 3 |
| I agree somewhat | 8 |
| I agree strongly | 9 |
| **General satisfaction** | |
| **I think MARIA can provide useful service for my medication adherence in the future, n** | |
| I disagree strongly | 0 |
| I disagree somewhat | 0 |
| I am neutral about it | 1 |
| I agree somewhat | 6 |
| I agree strongly | 13 |
| **Background history** | |
| **Disease, n** | |
| Ischemic dilated cardiomyopathy | 10 |
| Nonischemic dilated cardiomyopathy | 10 |
| **New York Heart Association, n** | |
| I | 16 |
| II | 4 |
| **Education level, n** | |
| Primary | 1 |
| Secondary | 9 |
| Higher level education/tertiary | 8 |
| Post degree | 2 |
| **Occupation, n** | |
| Unemployed or pensioner | 4 |
| Self-employed | 4 |
| Housewife | 3 |
| Engineer | 2 |
| Administrative | 4 |
| Teacher | 2 |
| Designer | 1 |
| **Ethnicity, n** | |
| Malay | 2 |
| Chinese | 2 |

| Demographic | Values |
|---|---|
| Bumiputra Sabah | 16 |

[a]MARIA: Medical Assistance and Rehabilitation Intelligent Agent.

## Discussion

### Principal Findings

The design thinking method provided an iterative process that actively engaged end users from the early stages of developing the MARIA prototype, a rule-based CA.

The involvement of a multidisciplinary group of stakeholders during the ideation phase facilitated the early conceptualization of the dialogue workflow, guided by psychoeducational theory—specifically, Adlerian therapy.

During the testing phase, the WoZ methodology and user satisfaction scoring were integrated into an observational study protocol. This approach enabled the collection of simulated real-world dialogues between patients and the MARIA prototype, operated by the wizard (pharmacists), allowing for iterative refinement and validation of the CA's conversational design.

The dialogues generated between the wizard (pharmacists) and the patients were systematically analyzed using coding analysis. This approach enabled the categorization of utterances into dialogue workflow components, speech acts, and topics, facilitating a structured evaluation of MARIA's conversational framework.

Speech acts—such as informing and expressing gratitude—were examined in relation to their associated topics and mapped to the dialogue workflow. This mapping validated the practical application of Adlerian theory, demonstrating its effectiveness in guiding the wizard to motivate patients toward medication adherence. Furthermore, the user satisfaction scores from patients confirmed the feasibility of applying Adlerian theory within the medication adherence dialogue workflow.

Additionally, the analysis identified instances where patient-initiated utterances—either new topics or responses—were not covered in the predefined dialogue workflow. These gaps highlighted areas for further refinement in MARIA's conversational design.

Building on this, the coding analysis reinforced the critical role of the wizard—played by an appropriate expert, in this case, pharmacists—as a key stakeholder in shaping how MARIA's dialogues should be modeled. For instance, it became evident that advising on medication interactions and side effects cannot be delegated to the CA, as these responses require human expertise to ensure patient safety. This insight guided the identification of various scenarios that must be accounted for from a patient safety perspective when designing MARIA's dialogue framework.

Furthermore, the user satisfaction scoring on human-likeness and trust highlighted the necessity of ensuring that MARIA's dialogues and use of emojis align with professional communication standards. Patients expressed a greater willingness to trust MARIA's advice on medication adherence when interactions were conducted professionally. This finding underscores the importance of designing CA interactions that balance humanlike engagement with a level of professionalism that fosters trust and credibility.

### Improvements

Through further analysis of the WoZ chatting history, we identified specific areas in MARIA's designed conversations that required optimization. These insights guided refinements to the current template design, ensuring a more effective and user-centered interaction experience. Based on these findings, we iterated on the conversational templates and provided the final version to the MARIA research team for future implementation.

In specific interactions, the MARIA medical team, drawing from their practical experience with patients with HF across various age groups, suggested that formal language use may be more suitable than casual language.

The use of words such as "cool" in MARIA's responses may create a more relaxed conversational style, which could be effective for younger patients but may not align with the preferences of older patients. Replacing "cool" with "excellent" could be more universally accepted across all age groups.

Specific messages should be designed to emphasize patients' responsibility in self-managing medication adherence. For example, MARIA should educate patients that they are not merely completing a task instructed by MARIA but actively working toward their own health goals. The messaging should reinforce that patients are empowered to take charge of their health, while MARIA serves as an assistant, supporting them in improving their health status.

Educating patients about medication in advance can help alleviate their concerns. MARIA should provide reference links to information on medication and HF for patients to review before following their medication plan. This approach can enhance patients' understanding of proper medication use, improve their awareness of potential side effects, and reduce the risk of misunderstandings about treatment effectiveness. Additionally, it may help prevent severe emergencies.

### Outcomes

The evaluation outcomes indicate that our conversational template design generally met stakeholders' needs. MARIA's natural language conversations, along with its encouragement and education strategies, are expected to support patients with HF in adhering to their medication. We identified several modifications that could enhance the applicability of the current conversational templates.

### Limitations

This section discusses the study's limitations and directions for future research. In this study, we were constrained by the

absence of a database containing basic medication knowledge and patient stories of successful adherence to HF medication at the prototype stage. Future development should focus on enriching MARIA's knowledge database to better support the designed education and encouragement strategies. The database should include comprehensive medication information from reliable sources and feature shared experiences of patients with HF who have successfully adhered to their treatment. Additionally, MARIA should be trained to provide tailored encouragement for patients facing various challenges in medication adherence. While linking to existing reputable HF associations worldwide is essential, collecting and curating real-life encouragement stories at the local level could improve cultural relevance and applicability. Furthermore, the study's participant pool was predominantly male, with limited female representation. This gender imbalance may affect the generalizability of the findings, and future research should ensure a more balanced representation to strengthen the applicability of the results.

Furthermore, as this is the initial stage of development, our focus was on covering a broad range of aspects rather than deeply exploring anthropomorphism. In future development stages, we plan to conduct a more detailed evaluation of anthropomorphism to enhance MARIA's human-like interactions.

## Conclusions

This study demonstrated that applying design thinking processes provides practical, interactive steps to engage users early in the design, prototyping, and testing of a CA for supporting patients in self-managing their medication. Furthermore, using the WoZ simulation method within an observational study protocol at the testing stage proved to be a valuable approach for refining the CA's interaction model, validating its functionality, and assessing system reliability, user expectations, and potential constraints. Results from the WoZ simulation and user satisfaction scores indicated that MARIA is a feasible and acceptable medication assistant CA. Additionally, patients expressed a general willingness to integrate MARIA into their daily routines to enhance medication adherence at home.

## Acknowledgments

Funding for the study was provided by Monash University Malaysia and the Persatuan Penyelidikan Klinikal Hospital Queen Elizabeth II. This study would not have been possible without the support of Monash University Malaysia students for the prototype research and development, and the Cardiology Department at Hospital Queen Elizabeth II, Sabah.

ChatGPT (OpenAI, Inc.) or other similar artificial intelligence tools were not used in the preparation of this manuscript.

## Data Availability

The manuscript provides the chat data. The MARIA prototype and the code are available for other researchers to use upon request. Researchers may also contact the main author to learn more about the design process.

## Authors' Contributions

NNA played a pivotal role in the study, handling conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, and resource management; also supervised the entire project, wrote the original draft, and participated in the review and editing of the manuscript. JT was involved in conceptualizing the study, performing formal analyses, and developing methodologies, contributing to the writing of the original draft and its subsequent review and editing. CHY contributed to conceptualization, data curation, formal analysis, funding acquisition, and methodology while also being involved in project administration, investigation activities, and the manuscript's review and editing process. HF developed the application software, ensuring its accurate description and integration within the research context, and also contributed to the literature review, as well as the review and editing of the manuscript. NFBK was involved in the investigation and project administration and participated in the review and editing of the manuscript, providing valuable feedback and insights. SBS contributed to the investigation and the manuscript's review and editing, ensuring the study's integrity and accuracy. VSY participated in the investigation and the review and editing process, providing critical revisions that enhanced the study's quality.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
User satisfaction of MARIA Interaction feedback. MARIA: Medical Assistance and Rehabilitation Intelligent Agent.
[PDF File (Adobe PDF File), 156 KB - cardio_v9i1e55846_app1.pdf ]

Multimedia Appendix 2
Samples of Wizard of Oz utterance analysis.
[PDF File (Adobe PDF File), 704 KB - cardio_v9i1e55846_app2.pdf ]

Multimedia Appendix 3
Excerpt from the coding analysis.
[DOCX File , 20 KB - cardio_v9i1e55846_app3.docx ]

## References


1. Savarese G, Lund LH. Global public health burden of heart failure. Card Fail Rev 2017 Apr;3(1):7-11 [FREE Full text] [doi: 10.15420/cfr.2016:25:2] [Medline: 28785469]
2. Lam C. Heart failure in Southeast Asia: facts and numbers. ESC Heart Fail 2015 Jun;2(2):46-49 [FREE Full text] [doi: 10.1002/ehf2.12036] [Medline: 28834655]
3. Bragazzi N, Zhong W, Shu J, Abu Much A, Lotan D, Grupper A, et al. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. Eur J Prev Cardiol 2021 Dec 29;28(15):1682-1690 [FREE Full text] [doi: 10.1093/eurjpc/zwaa147] [Medline: 33571994]
4. Management of Heart Failure. National Heart Association of Malaysia. 2019. URL: https://www.malaysianheart.org/publication/clinical-practice-guidelines/p/management-of-heart-failure-2019-4th-edition [accessed 2025-02-26]
5. Fitzgerald A, Powers JD, Ho PM, Maddox TM, Peterson PN, Allen LA, et al. Impact of medication nonadherence on hospitalizations and mortality in heart failure. J Card Fail 2011 Aug;17(8):664-669 [FREE Full text] [doi: 10.1016/j.cardfail.2011.04.011] [Medline: 21807328]
6. Ayaz T, Sahin SB, Sahin OZ, Bilir O, Rakıcı H. Factors affecting mortality in elderly patients hospitalized for nonmalignant reasons. J Aging Res 2014;2014:584315 [FREE Full text] [doi: 10.1155/2014/584315] [Medline: 25147737]
7. Kapoor J, Kapoor R, Ju C, Heidenreich PA, Eapen ZJ, Hernandez AF, et al. Precipitating clinical factors, heart failure characterization, and outcomes in patients hospitalized with heart failure with reduced, borderline, and preserved ejection fraction. JACC Heart Fail 2016 Jun;4(6):464-472 [FREE Full text] [doi: 10.1016/j.jchf.2016.02.017] [Medline: 27256749]
8. Shah D, Simms K, Barksdale D, Wu J. Improving medication adherence of patients with chronic heart failure: challenges and solutions. RRCC 2015 Jul:87 [FREE Full text] [doi: 10.2147/rrcc.s50658]
9. van der Wal MHL, Jaarsma T, van Veldhuisen DJ. Non-compliance in patients with heart failure; how can we manage it? Eur J Heart Fail 2005 Jan;7(1):5-17 [FREE Full text] [doi: 10.1016/j.ejheart.2004.04.007] [Medline: 15642526]
10. Ter Stal S, Sloots J, Ramlal A, Op den Akker H, Lenferink A, Tabak M. An embodied conversational agent in an eHealth self-management intervention for chronic obstructive pulmonary disease and chronic heart failure: exploratory study in a real-life setting. JMIR Hum Factors 2021 Nov 04;8(4):e24110 [FREE Full text] [doi: 10.2196/24110] [Medline: 34734824]
11. Swick RK. The accuracy of artificial intelligence (AI) chatbots in telemedicine. Journal of the South Carolina Academy of Science 2021;19(2):144-148 [FREE Full text]
12. Tudor Car L, Dhinagaran DA, Kyaw BM, Kowatsch T, Joty S, Theng Y, et al. Conversational agents in health care: scoping review and conceptual analysis. J Med Internet Res 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: 10.2196/17158] [Medline: 32763886]
13. Watts R, Pietrzak D. Adlerian “encouragement” and the therapeutic process of solution‐focused brief therapy. Journal of Counseling & Development 2011 Dec 23;78(4):442-447 [FREE Full text] [doi: 10.1002/j.1556-6676.2000.tb01927.x]
14. Carlson J, Watts RE, Maniacci M. Adlerian Therapy: Theory and Practice. Washington, DC: American Psychological Association; 2006.
15. DuPont D. ChoosingTherapy. 2023 Aug 13. URL: https://www.choosingtherapy.com/adlerian-therapy/ [accessed 2025-02-27]
16. Wong Y. The psychology of encouragement. The Counseling Psychologist 2014 Oct 27;43(2):178-216 [FREE Full text] [doi: 10.1177/0011000014545091]
17. Pangaro P. Questions for conversation theory. Paul Pangaro (blog page). 2016. URL: https://pangaro.com/published/Pangaro-ASC_2016-Conversation_Theory_in_One_Hour-Final_Draft.pdf [accessed 2024-08-20]
18. Spangenberg E, Shavitt S, Brock TC. Persuasion: psychological insights and perspectives. Journal of Marketing Research 1996 May;33(2):259. [doi: 10.2307/3152156]
19. Agarwal R, Wadhwa M. Review of state-of-the-art design techniques for chatbots. SN Comput Sci 2020 Jul 29;1(5):246. [doi: 10.1007/s42979-020-00255-3]
20. Dhinagaran DA, Martinengo L, Ho MR, Joty S, Kowatsch T, Atun R, et al. Designing, developing, evaluating, and implementing a smartphone-delivered, rule-based conversational agent (DISCOVER): development of a conceptual framework. JMIR Mhealth Uhealth 2022 Oct 04;10(10):e38740 [FREE Full text] [doi: 10.2196/38740] [Medline: 36194462]
21. Chatbots reset: a framework for governing responsible use of conversational AI in health care. World Economic Forum. URL: https://www.weforum.org/reports/chatbots-reset-a-framework-for-governing-responsible-use-of-conversational-ai-in-health [accessed 2025-02-25]
22. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30 [FREE Full text] [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
23. Ujiro T, Tanaka H, Adachi H, Kazui H, Ikeda M, Kudo T, et al. Detection of dementia from responses to atypical questions asked by embodied conversational agents. Proc Interspeech 2018:1691-1695 [FREE Full text] [doi: 10.21437/interspeech.2018-1514]

24. Gross C, Schachner T, Hasl A, Kohlbrenner D, Clarenbach CF, Wangenheim FV, et al. Personalization of conversational agent-patient interaction styles for chronic disease management: two consecutive cross-sectional questionnaire studies. J Med Internet Res 2021 May 26;23(5):e26643 [FREE Full text] [doi: 10.2196/26643] [Medline: 33913814]
25. Ferguson G, Allen J, Galescu L, Quinn J, Swift M. Cardiac: an intelligent conversational assistant for chronic heart failure patient heath monitoring. Washington, DC: Association for the Advancement of Artificial Intelligence; 2009 Presented at: AAAI Fall Symposium; November 4-7, 2009; Arlington, VA URL: https://cdn.aaai.org/ocs/915/915-4134-1-PB.pdf
26. Moulik S, Chatterjee S. Dil - a conversational agent for heart failure patients. In: Proceedings of the 54th Hawaii International Conference on System Sciences. 2021 Presented at: The 54th Hawaii International Conference on System Sciences; January 5-8, 2021; Kauai, HI p. 3546-3555 URL: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/51624983-43eb-4a20-8eb6-c5b6f15eaee3/content [doi: 10.24251/hicss.2021.430]
27. Lobo J, Ferreira L, Ferreira AJ. CARMIE: a conversational medication assistant for heart failure. International Journal of E-Health and Medical Communications (IJEHMC) 2017;8(4):21-37. [doi: 10.4018/ijehmc.2017100102]
28. Apergi LA, Bjarnadottir MV, Baras JS, Golden BL, Anderson KM, Chou J, et al. Voice interface technology adoption by patients with heart failure: pilot comparison study. JMIR Mhealth Uhealth 2021 Apr 01;9(4):e24646 [FREE Full text] [doi: 10.2196/24646] [Medline: 33792556]
29. Dingler T, Kwasnicka D, Wei J, Gong E, Oldenburg B. The use and promise of conversational agents in digital health. Yearb Med Inform 2021 Aug 03;30(1):191-199 [FREE Full text] [doi: 10.1055/s-0041-1726510] [Medline: 34479391]
30. Schachner T, Gross C, Hasl A, V Wangenheim F, Kowatsch T. Deliberative and paternalistic interaction styles for conversational agents in digital health: procedure and validation through a web-based experiment. J Med Internet Res 2021 Jan 29;23(1):e22919 [FREE Full text] [doi: 10.2196/22919] [Medline: 33512328]
31. Brinke JST, Fitte C, Anton E, Meier P, Teuteberg F. With a little help from my conversational agent: towards a voice assistant for improved patient compliance and medication therapy safety. 2021 Presented at: The 14th International Joint Conference on Biomedical Engineering Systems and Technologies; February 11-13, 2021; Virtual URL: https://tinyurl.com/yhu8vhyy [doi: 10.5220/0010411707890800]
32. Adler A. The Practice and Theory of Individual Psychology (Vol. 133). London, UK: Routledge; 2013.
33. Abdullah N, Grundy J, McIntosh J, How YC, Saharuddin S, Tat KK, et al. Using work system design, user stories and emotional goal modeling for an mhealth system. 2020 Presented at: 2020 IEEE First International Workshop on Requirements Engineering for Well-Being, Aging, and Health (REWBAH); August 31, 2020; Zurich, Switzerland. [doi: 10.1109/rewbah51211.2020.00007]
34. Social cognitive theory. Rural Health Information Hub. 2021. URL: https://www.ruralhealthinfo.org/toolkits/health-promotion/2/theories-and-models/social-cognitive [accessed 2025-03-18]
35. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
36. Stajkovic A, Sergent K. Social cognitive theory. In: Management. Oxford, UK: Oxford University Press; 2019:1-14.
37. Nabavi RT, Bijandi MS. ResearchGate. 2012 Jan. URL: https://www.researchgate.net/publication/267750204_Bandura's_Social_Learning_Theory_Social_Cognitive_Learning_Theory [accessed 2025-03-19]
38. Guéguen N, Joule RV, Courbet D, Halimi-Falkowicz S, Marchand M. Repeating “yes” in a first request and compliance with a later request: the four walls technique. Social Behavior and Personality 2013 Mar;41(2):199-202 [FREE Full text] [doi: 10.2224/sbp.2013.41.2.199]
39. Cialdini R, Vincent J, Lewis S, Catalan J, Wheeler D, Darby B. Reciprocal concessions procedure for inducing compliance: the door-in-the-face technique. Journal of Personality and Social Psychology 1975 Feb;31(2):206-215 [FREE Full text] [doi: 10.1037/h0076284]
40. Fetrati H, Chan G, Orji R. Chatbots for sexual health improvement: a systematic review. International Journal of Human–Computer Interaction 2024 Jul 15;41(4):1997-2019. [doi: 10.1080/10447318.2024.2376808]
41. Liu W, Jiang M, Li W, Mou J. How does the anthropomorphism of AI chatbots facilitate users' reuse intention in online health consultation services? The moderating role of disease severity. Technological Forecasting and Social Change 2024 Jun;203:123407. [doi: 10.1016/j.techfore.2024.123407]
42. Nkabane-Nkholongo E, Mokgatle M, Bickmore T, Julce C, Jack BW. Adaptation of the Gabby conversational agent system to improve the sexual and reproductive health of young women in Lesotho. Front Digit Health 2023 Oct 4;5:1224429. [doi: 10.3389/fdgth.2023.1224429] [Medline: 37860039]
43. Candello H, Pinhanez C. Designing conversational interfaces. In: Tutorial. XV Simpósio Sobre Fatores Humanos em Sistemas Computacionais. São Paulo, Brazil: BDBComp; Oct 2016:1-33.
44. Shumanov M, Johnson L. Making conversations with chatbots more personalized. Computers in Human Behavior 2021 Apr;117:106627 [FREE Full text] [doi: 10.1016/j.chb.2020.106627]
45. Stark L, Crawford K. The conservatism of emoji: work, affect, and communication. Social Media + Society 2015 Oct 08;1(2):205630511560485 [FREE Full text] [doi: 10.1177/2056305115604853]
46. James S. Stream Creative. 2020. URL: https://www.streamcreative.com/chatbot-scripts-examples-templates [accessed 2025-02-27]

47. Fadhil A, Schiavo G. Designing for health chatbots. arXiv. Preprint posted online on February 24, 2019 URL: https://arxiv.org/abs/1902.09022 [accessed 2024-08-08]
48. Zhang J, Oh YJ, Lange P, Yu Z, Fukuoka Y. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: viewpoint. J Med Internet Res 2020 Sep 30;22(9):e22845 [FREE Full text] [doi: 10.2196/22845] [Medline: 32996892]
49. Caplan JE, Adams K, Boyd RL. Personality and language. In: The Wiley Encyclopedia of Personality and Individual Differences. Hoboken, NJ: Wiley; 2000:311-316.
50. Janson A. How to leverage anthropomorphism for chatbot service interfaces: the interplay of communication style and personification. Computers in Human Behavior 2023 Dec;149:107954. [doi: 10.1016/j.chb.2023.107954]
51. Endriss U, Maudet N, Sadri F, Toni F. Logic-based agent communication protocols. In: Dignum F, editor. Advances in Agent Communication (ACL 2003). Berlin/Heidelberg, Germany: Springer; 2004:91-107.
52. Okamoto M, Yang Y, Ishida T. Wizard of Oz method for learning dialog agents. In: Klusch M, Zambonelli F, editors. Cooperative Information Agents V. CIA 2001. Berlin, Heidelberg, Germany: Springer; 2001:20-25.
53. Schlögl S, Doherty G, Luz S. Wizard of Oz experimentation for language technology applications: challenges and tools. Interact Comput 2014 May 09;27(6):592-615. [doi: 10.1093/iwc/iwu016]
54. Kelley J. An empirical methodology for writing user-friendly natural language computer applications. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. New York, NY: Association for Computing Machinery; 1983 Presented at: The SIGCHI conference on Human Factors in Computing Systems; December 12-15, 1983; Boston, MA p. 193-196. [doi: 10.1145/800045.801609]
55. Fraser NM, Gilbert G. Simulating speech systems. Computer Speech & Language 1991 Jan;5(1):81-99. [doi: 10.1016/0885-2308(91)90019-m]
56. Hauptmann A. Speech and gestures for graphic image manipulation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1989 Jan 01 Presented at: The SIGCHI Conference on Human Factors in Computing Systems; April 30 to June 4, 1989; Austin, TX p. 241-245. [doi: 10.1145/67449.67496]
57. De Marconnay P, Crowley J, Salber D. Visual interpretation of faces in the NEIMO multimodal test-bed. 1993 Presented at: IJCAI 93 Conference; San Francisco, CA; August 28 to September 3, 1993.
58. Salber D, Coutaz J. A Wizard of Oz platform for the study of multimodal systems. 1993 Jan 01 Presented at: INTERCHI93: Conference on Human Factors in Computing; April 24-29, 1993; Amsterdam, The Netherlands p. 95-96. [doi: 10.1145/259964.260126]
59. Eberhart Z, Bansal A, McMillan C. A Wizard of Oz study simulating API usage dialogues with a virtual assistant. IIEEE Trans Software Eng 2022 Jun 1;48(6):1883-1904. [doi: 10.1109/tse.2020.3040935]
60. Pink S, Osz K, Fors V, Lanzeni D. Simulating and trusting in automated futures: anthropology and the Wizard of Oz. In: Kazubowski-Houston M, Auslander M, editors. In Search of Lost Futures. Cham, Switzerland: Palgrave Macmillan; Feb 17, 2021:195-218.
61. Poisson C, Barré J, Bourmaud G, Forzy JF. Driver behavior in conditional automation: comparison of driving simulator and Wizard of Oz conditions. 2020 Presented at: CHI '20: CHI Conference on Human Factors in Computing Systems; April 25-30, 2020; Honolulu, HI p. 1-7. [doi: 10.1145/3334480.3382854]
62. Manstetten D, Marberger C, Beruscha F. Wizard-of-Oz Experiments in Real Traffic - Can They Restart Human Factors?. Darmstadt, Germany: Technical University of Darmstadt; 2019:21-31.
63. Bellucci A, Zarraonandia T, Díaz P, Aedo I. Welicit: a wizard of Oz tool for VR elicitation studies. 2021 Presented at: Human-Computer Interaction – INTERACT 2021: 18th IFIP TC 13 International Conference; August 30 to September 3, 2021; Bari, Italy p. 82-91. [doi: 10.1007/978-3-030-85607-6_6]
64. Torstensson N, Susi T, Wilhelmsson U, Lebram M. Wizard of Oz and the design of a multi-player mixed reality game. 2020 Presented at: HCI in Games: Second International Conference, HCI-Games 2020, Held as Part of the 22nd HCI International Conference, HCII 2020; July 19-24, 2020; Copenhagen, Denmark p. 218-232. [doi: 10.1007/978-3-030-50164-8_15]
65. Dam RF, Siang TY. 10 insightful design thinking frameworks: a quick overview. Interaction Design Foundation. 2020. URL: https://www.interaction-design.org/literature/article/design-thinking-a-quick-overview [accessed 2025-03-19]
66. Steinke G, Al-Deen MS, LaBrie RC. Innovating information system development methodologies with design thinking. 2017 Presented at: The 5th International Conference on Applied Innovations in IT (ICAIIT); March 16, 2017; Koethen, Germany.
67. Plattner H, Meinel C, Leifer LJ, editors. Design Thinking: Understand, Improve, Apply. Understanding Innovation. Berlin/Heidelberg, Germany: Springer-Verlag; 2011:xiv-xvi.
68. Fetrati H. (. Malaysia: Unpublished; 2021:A.
69. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
70. Walker S. Types of transcription: understanding different transcription services. New Media Services. 2018 Jun 1. URL: https://newmediaservices.com.au/types-of-transcription/ [accessed 2024-08-08]

71. Meena HK, Saha I, Mondal KK, Prabhakar TV. An approach to workflow modeling and analysis. In: Proceedings of the 2005 OOPSLA Workshop on Eclipse Technology EXchange - eclipse '05. New York, NY: Association for Computing Machinery; 2005 Presented at: eclipse '05: Proceedings of the 2005 OOPSLA workshop on Eclipse technology eXchange; October 16-17, 2005; San Diego, CA p. 85-89. [doi: 10.1145/1117696.1117714]
72. Song Y, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 2015 Apr 25;27(2):130-135 [FREE Full text] [doi: 10.11919/j.issn.1002-0829.215044] [Medline: 26120265]
73. Bonial C, Marge M, Foots A, Gervits F, Hayes C, Henry C, et al. Laying down the yellow brick road: Development of a wizard-of-oz interface for collecting human-robot dialogue. arXiv Preprint posted online on October 17, 2017.
74. Nielsen J, Norman D. Why you only need to test with 5 users. Nielsen Norman Group. 2000. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2025-03-19]
75. Mich L, Sakhnini V, Berry D. To group or not to group? Group sizes for requirements elicitation. Information and Software Technology 2023 Aug;160:107229. [doi: 10.1016/j.infsof.2023.107229]
76. Boddy C. Sample size for qualitative research. QMR 2016 Sep 12;19(4):426-432. [doi: 10.1108/qmr-06-2016-0053]
77. The Pharmacy Practice and Development Division. The Ministry of Health Malaysia. URL: https://pharmacy.moh.gov.my/en/content/pharmacy-practice-development-division.html [accessed 2025-03-18]
78. Protocol medication therapy adherence clinic: heart failure. Official Portal of Pharmaceutical Services Programme. Selangor, Malaysia: Ministry of Health Malaysia URL: https://www.pharmacy.gov.my/v2/en/documents/protocol-medication-therapy-adherence-clinic-heart-failure.html [accessed 2024-08-08]
79. Vanderveken D. Meaning and Speech Acts: Principles of Language Use (Vol. 1). Cambridge, UK: Cambridge University Press; 1990.
80. Hoffman R, Mueller ST, Klein G, Litman J. Metrics for explainable AI: challenges and prospects. arXiv Preprint posted online on February 1, 2019 [FREE Full text]


## Abbreviations

**AI:** artificial intelligence
**CA:** conversational agent
**CRC:** Clinical Research Center
**DH:** digital health
**HF:** heart failure
**MARIA:** Medical Assistance and Rehabilitation Intelligent Agent
**WoZ:** Wizard of Oz


*Edited by A Coristine; submitted 27.12.23; peer-reviewed by S Staves, L Moradbakhti; comments to author 19.06.24; revised version received 09.09.24; accepted 09.12.24; published 10.04.25.*

*Please cite as:*
*Abdullah NN, Tang J, Fetrati H, Kaukiah NFB, Saharudin SB, Yong VS, Yen CH*
*MARIA (Medical Assistance and Rehabilitation Intelligent Agent) for Medication Adherence in Patients With Heart Failure: Empirical Results From a Wizard of Oz Systematic Conversational Agent Design Clinical Protocol*
*JMIR Cardio 2025;9:e55846*
*URL: https://cardio.jmir.org/2025/1/e55846*
*doi:10.2196/55846*
*PMID:*

# Evaluation of a Virtual Home Health Heart Failure Program: Mixed Methods Study

Nilufeur McKay[1], DNP; Rosemary Saunders[1], PhD; Helene Metcalfe[1], PhD; Sue Robinson[1], MSc; Peter Palamara[1], MSc; Kellie Steer[2], BSc, MBA; Jeannie Yoo[3,4], MA, MBBS; Miles Ranogajec[5], BSc, MPhysio; Lisa Whitehead[1], PhD; Beverley Ewens[1], PhD

[1]School of Nursing and Midwifery, Edith Cowan University, 270 Joondalup Drive, Joondalup, Australia

[2]Ramsay Connect, Perth, Australia

[3]Australian Unity, Sydney, Australia

[4]College of Medicine and Public Health, Flinders University, Bedford Park, Australia

[5]Ramsay Connect, Geelong, Australia

**Corresponding Author:**
Nilufeur McKay, DNP
School of Nursing and Midwifery, Edith Cowan University, 270 Joondalup Drive, Joondalup, Australia

## Abstract

**Background:** Heart failure is a prevalent and debilitating condition, affecting millions globally and imposing a significant burden on patients, families, and health care systems. Despite advancements in medical treatments, the gap in effective, continuous, and personalized supportive care remains glaringly evident. To address this pressing issue, virtual health care services delivered by interdisciplinary teams represent a promising solution. Understanding the outcomes and experience of remote monitoring–enabled interdisciplinary chronic disease management programs can inform resource allocation and health care policy decisions.

**Objective:** The purpose of this study was to evaluate the clinical and behavioral outcomes of patients undertaking a Virtual Home Health Heart Failure Program (VHHHFP) and explore the experiences of patients and health care practitioners (HCPs).

**Methods:** The VHHHFP is a virtual postdischarge support service for patients with heart failure that includes an intensive 3-month period followed by a maintenance period delivered by an interdisciplinary team. A mixed methods study was conducted with patients and HCPs. Self-reported outcome data (KCCQ-12 [Kansas City Cardiomyopathy Questionnaire-12], PHQ-4 [Patient Health Questionnaire-4], PAM-13 [Patient Activation Measure-13], and PREMs [Patient Reported Experience Measures]) were obtained from the records of patients (N=49) who completed the intensive phase of the VHHHFP, and interviews were conducted with patients (n=9) and HCPs (n=6). A paired *t* test was used to compare quantitative data before and after the 3-month intervention, and a thematic qualitative analysis was undertaken of interview data.

**Results:** Thirty-one of the 55 (77.5%) patients completed the baseline and 3-month follow-up KCCQ-12 assessment. The mean KCCQ-12 summary score at 3 months was 72.20 (SD 20.2), which was significantly higher than the mean summary score at baseline of 50.51 (SD 17.59; *P*<.001). These findings were similar for the KCCCQ-12 subscales: physical limitations (mean 47.09, SD 29.7 and mean 69.43, SD 22.6; *P*<.001), quality of life (mean 43.75, SD 21.7 and mean 62.91, SD 25.7; *P*<.001), symptom frequency (median 60.40, IQR 1-100 and median 91.70, IQR 35.40; *P*<.001), and social limitation (median 50.0, IQR 1-100 and median 82.50, IQR 32.50; *P*<.001). The PHQ-4 measure of psychological health was completed by 32 (80%) patients. The median scores at baseline and follow-up for total distress (median 1.50, IQR 0-7 and median 0.0, IQR 0-8; *P*<.02), and the anxiety subscale (median 1.0, IQR 0-6 and median 0.0, IQR 0-4; *P*<.02) reduced over time. Six hospital admissions were recorded (10.2% of 49 patients) within 30 days. Nine patient interviews aligned with the value-based health care (VBHC) Capability, Comfort, and Calm (CCC) framework. Three themes were identified, which are as follows: (1) enhanced patient capability, (2) improved patient comfort, and (3) positive influences on calm. Six health care professionals shared experiences of the VHHHFP, with three emerging themes: (1) improved patient capability through shared decision-making, (2) improving capability through care practices, and (3) promoting comfort and calm through virtual coordination and collaboration.

**Conclusions:** The use of technologies to support the management of HF is an area of growth. This study contributes to the understanding of how remote patient monitoring with interdisciplinary chronic disease support, integrated into an existing system, can improve clinical outcomes for patients.

***(JMIR Cardio 2025;9:e64877)*** doi:10.2196/64877

**KEYWORDS**

heart failure; patient care team; telemedicine; virtual health; mixed methods study; healthcare systems; supportive care; virtual healthcare; monitoring support program; quality of life; Australian; value-based healthcare

## Introduction

Heart failure (HF) has been acknowledged as one of the Western world's most significant public health issues [1]. This chronic condition results in reduced quality of life, creating a burden for health care systems in terms of resource use and financial cost [2-4]. Globally, HF is described as an epidemic, affecting more than 64 million people worldwide [5], and a diagnosis of HF is associated with high rates of morbidity and mortality, particularly in low- and middle-income countries [6]. The percentage of the population diagnosed with HF around the world varies among populations [7], but globally is estimated to be between 1% and 3% of the total population [8]. In Australia, it is estimated that 1% - 2% of the population are diagnosed with HF, compared to a prevalence of 2.4% - 3% in the United States [2] and between 1.3% and 6.7% in Asia [3]. In 2017 - 2018, an estimated 102,000 (0.5%) people self-reported living with HF within Australia, with around 179,000 hospitalizations in 2020 - 2021 attributed to HF or cardiomyopathy as the primary diagnosis [9]. The prevalence of HF is predicted to increase due to the aging population, improved treatment of acute cardiac events, and availability of evidence-based therapies for those with HF [2]. It is estimated that by 2023, cases of HF in Australia will increase to 750,000 [10]. The majority of health care costs for people with HF are associated with an increasing rate of hospitalizations due to poor self-care, nonadherence to treatment, or inability to access medications [4]. Research suggests that most patients (80%) living with HF are reliant on their general practitioner (GP) for ongoing management and support [11].

A growing body of evidence supports the use of digital health technology in improving patient outcomes [12-14], with telemonitoring [15] and digital health becoming central to health care [16]. Virtual health care has become an indispensable component of contemporary care delivery, which enables those with chronic conditions to stay connected to online supportive environments and clinicians to establish two-way communication and noninvasive monitoring for patients in remote locations [17]. The COVID-19 pandemic expedited the adoption of telehealth globally. However, the evaluation of telehealth outcomes has not necessarily matched the pace of its uptake [18]. In response, there has been an increase in the exploration of remote and virtual patient monitoring and care models to manage and improve the outcomes of patients with HF [12]. However, the use of virtual HF programs remains in its infancy in Australia. A recent systematic review identified that telemonitoring, remote patient management, and patient self-empowerment as an integrated approach performed best in terms of readmission rates and overall hospital visits [17], with further research into this approach needed [19]. This study aimed to evaluate the clinical and behavioral outcomes of patients undertaking a Virtual Home Health Heart Failure Program (VHHHFP) and to explore the experiences of patients and health care practitioners (HCPs) who participated in the program.

## Methods

### Research Design and Study Population

A mixed methods study was conducted in collaboration with patients and HCPs. Self-reported outcome data (KCCQ-12 [Kansas City Cardiomyopathy Questionnaire-12], PHQ-4 [Patient Health Questionnaire-4], PAM-13 [Patient Activation Measure-13], and PREMs [Patient Reported Experience Measures]) were obtained from the records of 55 patients who completed the intensive phase (0 - 3 months) of the VHHHFP. Interviews were conducted with 9 patients and 6 HCPs. A mixed methods approach was selected as it provides the opportunity to integrate quantitative findings in the analysis of the qualitative data [20]. This is particularly important in the evaluation of a chronic disease management program, where there is a need to understand from a patient's perspective how programs impact or fail to impact health outcomes observed from a purely quantitative approach. This study method provides a richer level of understanding of content, processes, and policies within programs [21]. The study is reported in accordance with the mixed methods reporting guidelines by Lee et al [22]. The study population comprised patients participating in the intensive phase of the program, the clinicians delivering the program, and other clinicians external to the program but involved in the care of the patients.

### The VHHHFP

The VHHHFP is a virtually delivered postdischarge support service for patients with HF. The program aims to (1) improve HF symptoms, quality of life, and physical and social limitations, (2) improve HF self-management skills and capabilities, (3) improve patient understanding of medications and therapy adherence, (4) reduce signs and symptoms of anxiety and depression associated with HF, and (5) reduce preventable hospital admissions through collaborative care practices.

Suitable patients who meet the inclusion criteria (Textbox 1) can be referred to the program by a member of their inpatient care team during a hospital admission. Services are delivered via telehealth by a clinical nurse specialist (CNS), a clinical nurse or registered nurse, a dietitian, and a physiotherapist. Interventions provided include care coordination, remote patient monitoring of vital signs and symptoms, nurse-led medication titration (as directed by the patient's cardiologist or GP), education, virtually delivered exercise (when clinically appropriate), and support for nutrition and weight management. The necessary equipment is provided to patients at no cost. The program operates during business hours, with an HF action plan provided to patients for out-of-hours concerns. The program integrates with primary and specialist care teams, with medical governance either provided by the patient's existing cardiologist or GP, as per the specialist's preference. The program's intensive phase is delivered over a period of 3 months and includes an initial visit by one of the VHHHFP clinical nurses, who

described the program to the patient and obtained consent to participate in the program. Initial screening surveys were completed, and the patient received a comprehensive description of the home monitoring equipment. Upon discharge, the patient received intensive care coordination by an interdisciplinary team with remote monitoring, medication titration, and self-care management support. More details of these individual components are represented in Figure 1. Following this, the patient enters a 3 - 12-month maintenance phase based on the patient's needs to embed long-term self-management behaviors. The maintenance phase of the program was not included in this study. Upon completion, the patient is discharged to their primary care clinician with an ongoing plan of care for their HF. At this point, the patient does not have further access to the VHHHFP.

Several validated self-reporting assessment tools were used to monitor the patient's progress while on the program. The measures were selected to be consistent with the health outcome measures for patients with HF recommended by the International Consortium for Health Outcomes [23]. Self-reported outcomes included the KCCQ-12 [24], the PHQ-4 for anxiety and depression [25], and the PAM-13 [26] for assessing engagement with health care and self-management, and a PREM survey. The PREM survey was adapted and modified from the Australian Hospital Patient Experience Question Set [27] to suit the context of the VHHHFP.

**Textbox 1.** Virtual Home Health Heart Failure Program inclusion and exclusion criteria.

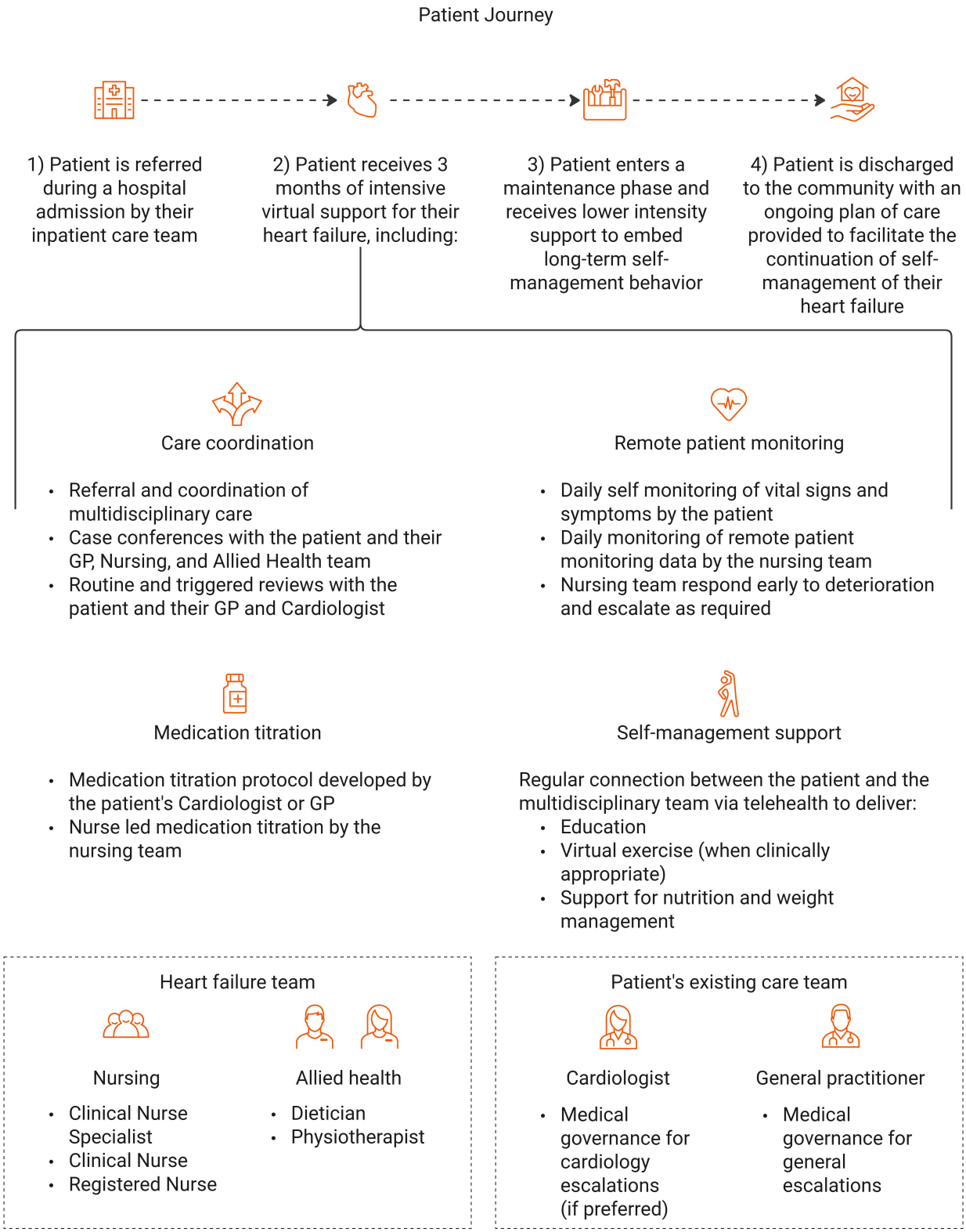
**Patient inclusion criteria**

- New or existing diagnosis of heart failure (heart failure with reduced ejection fraction (HFrEF) or heart failure with preserved ejection fraction (HFpEF).
- New York Heart Association (NYHA) functional class I, II, or III.
- Patient referral to the Virtual Home Health Heart Failure Program (VHHHFP) from a clinician at one of the three hospital sites.

**Exclusion criteria**

- Discharged to a residential aged care facility.
- NYHA functional class IV.
- Hemodynamic instability.
- Recurrent arrhythmias or unstable angina requiring investigation.
- Deteriorating renal function, defined as estimated glomerular filtration rate (eGFR) <30 mL/min.
- Continuous oxygen requirement at rest.
- Existing cognitive impairment.

**Figure 1.** The patient journey through the Virtual Home Health Heart Failure Program (a 3-months intensive phase followed by a maintenance period). GP: general practitioner.



## Recruitment and Consent

### *Patient Participants*

Patients provided consent for their quantitative data to be collected for evaluation on commencement of the program. Patients enrolled in the intensive phase of the VHHHFP were sent an email by the HF CNS, informing them about the study and the process to participate in an interview. Patients who were interested in participating in the semistructured interview contacted the study research assistant via email and were given the opportunity to ask any questions they had. Written informed consent was obtained before the commencement of the interviews.

### *Health Care Practitioner Participants*

HCPs (including HF CNSs, physiotherapists, GPs, and cardiologists) were eligible to participate if they had delivered care to the patients within the 3-month intensive phase of the program. Health care participants involved in the care of the recruited patient participants were sent an email to inform them about the study, and who then contacted the research team directly. HCP participants provided written consent before the commencement of interviews.

### *Setting*

Three hospital sites across Australia, which varied in size from 250 to 900 beds.

### *Data Collection*

Data were collected via surveys and individual semistructured interviews. The quantitative data were collected as part of the VHHHFP, and the qualitative data were collected to explore patients' and stakeholders' perspectives of the program.

### *Quantitative Data Collection*

Quantitative data included baseline demographic information and patient self-reported outcome data via the home digital platform as part of the VHHHFP. Baseline patient demographic and clinical characteristics collected included age, identified gender, HF subtype (heart failure with reduced ejection fraction or heart failure with preserved ejection fraction), NYHA Functional Class, and the Length of stay, Acuity on admission, Comorbidity, and Emergency room usage (LACE) index as a measure of patients' risk of 30-day readmission [28]. This information was extracted by a program administrator and provided to the research team. Patients completed the self-reported outcomes directly on the home digital platform at various times during the intensive phase. Pretest data (baseline) and posttest data (3 months after enrollment) were collected. Patients were followed up at key time points, including enrollment (baseline) and at 12 weeks, to measure health status (KCCQ-12), PREMs, anxiety and depression (PHQ-4), and engagement with health care and self-management (PAM-13). Data submitted up to 14 days before or post the expected completion date (70 - 98 d) of the 12-week program were included for analysis. A secure messaging platform was used to transfer patient data to the research team. Data did not contain identifiable patient data and were grouped via an ID number for analysis on a per-patient basis. For those participants who took part in the qualitative phase of the study, IDs were linkable via a separately stored key that provided patient contact details.

### *Qualitative Data Collection*

Individual interviews were conducted to obtain information about the patient and clinician perspectives of the program. The value-based health care (VBHC) Capability, Comfort, and Calm (CCC) framework [29] guided the development of the patient and HCP interview questions (Multimedia Appendix 1). Patients were given a choice to complete interviews in person or virtually. The interviews were audio-recorded and later transcribed by a member of the research team. Patients were given the option to be interviewed individually or with a family member, and permission was sought to audio record the interview. One-to-one individual interviews with health care providers were undertaken by a member of the research team who did not have a professional relationship with the participants. Health care provider participants responded to 8 open-ended questions (Multimedia Appendix 1). Interviews were audio-recorded and were up to an hour in duration.

### *Quantitative Analysis*

All data were exported from Microsoft CSV files to SPSS (version 28.0.1; IBM Corp) for analysis. Frequency counts were used to describe data pertaining to patient demographics and clinical status, program experience, and hospital readmissions. The normality of the data on patients' physical and psychological health status and health care behaviors at each time point was assessed graphically and numerically using the Shapiro-Wilk Test for small sample sizes. For data that was normally distributed, the mean and SD were used as measures of central tendency and dispersion, and dependent-samples *t* tests were used to test for changes over time (Time 1 and Time 2) in patient-reported outcomes. For data that were nonnormally distributed, the median and IQR, and the Wilcoxon signed-rank test for dependent samples were used. Frequency counts were also provided for patient demographic and clinical status data. Patient datasets that had missing pre- or postsurvey data were excluded from the analysis. Patients who died during the study were excluded as required by the organization's ethics committee.

### *Qualitative Analysis*

Thematic analysis was then undertaken using Braun and Clarke's six-step approach (familiarization with the data, generating codes, constructing themes, revising and defining themes, and producing the report) [30] to identify patterns of meaning to understand the patient and HCPs' experiences of the VHHHFP. The audio recordings were transcribed verbatim and verified against the recording, and the transcripts were read several times before initial coding by one researcher (HM) and then independently by two researchers (RS and NM). Following initial coding, the researchers discussed the coding and reached consensus on categories. Data were organized and managed using Microsoft Excel. Research team discussions then informed the development of themes and subthemes using the VBHC framework [31], comprising the concepts of CCC.

To ensure qualitative rigor and trustworthiness, data were collected until data saturation was achieved, which was based

on the appraisal of the data collected and the rich dialogue that related to the study aim. Member checking was not possible due to the single participant interaction. Discussion and interpretations of the data ensured credibility. The presentation of the findings will guide other researchers in the transferability of the findings. The researchers maintained a clear record of the reflexive analysis process, supporting the confirmability of the findings.

### Ethical Considerations

The research study was conducted following the Australian National Statement on Ethical Conduct of Human Research (2023), developed by the National Health and Medical Research Council, the Australian Research Council, and Universities Australia. Permission to conduct the study was sought and approved by an Australian institutional review board, the Ramsay Health Care Human Research Ethics Committee (approval no 2022/PID/2031), and the Edith Cowan University Human Research Ethics Committee (approval no 2022- ‐ 03864). Research procedures were followed under the ethical standards of the approving national Human Research Ethics Committees and with the WMA Declaration of Helsinki. Patients provided informed consent for their quantitative data to be collected for evaluation at the commencement of the program. Patients enrolled in the intensive phase of the VHHHFP were invited to participate in an interview via an email from the HF CNS, which informed them about the study and the process to participate in an interview. Patients and staff interested in participating in the qualitative component of the study were required to contact the study research assistant via email. All participants were provided with a participant information form, which provided detailed information about the study and what their involvement comprised. Potential participants were also informed they would receive a US $35 gift voucher as acknowledgment of their semistructured interview participation. Written informed consent was obtained before the commencement of all interviews. Participants were informed that their participation was voluntary and that they were free to withdraw from the study at any time; however, following data analysis, their data could not be removed. No personal identifying information was collected.

## Results

### Baseline Characteristics

From June 1, 2022, to November 30, 2022, a total of 55 patients were enrolled into the VHHHFP across 3 Australian hospital sites. Data for 5 patients were not reported in line with the ethical approval granted: 3 patients died during the intensive phase of the program, and 2 died during the data collection period following completion of the intensive phase of the program. In addition, 1 patient who enrolled but did not complete any questionnaire or biometric information was excluded. The demographic and clinical characteristics of the remaining 49 patients who enrolled and engaged with the program are presented in Table 1. Of these 49 patients, 7 withdrew and did not complete the program: 3 patients due to hospital admission, and 2 patients failed to complete the intensive phase of the program within the nominated time (98 days from enrollment data). Data for these 7 patients related to changes in physical and psychological health and health care behavior, as well as their experiences with the program, were subsequently excluded from the analysis.

**Table .** Demographic and clinical characteristics of patients enrolled and engaged in the Virtual Home Health Heart Failure Program.

| Variable | Patient, n (%) |
|---|---|
| Sex (N=49) | |
| Male | 31 (63) |
| Female | 18 (37) |
| Age (years; N=49) | |
| 41 - 50 | 3 (6) |
| 51 - 60 | 4 (8) |
| 61 - 70 | 17 (35) |
| 71 - 80 | 17 (35) |
| 81 - 90 | 8 (16) |
| NHYA[a] classification (N=42) | |
| I | 0 (0) |
| II | 34 (81) |
| III | 7 (17) |
| IV | 1[b] (2) |
| Ejection fraction (%; N=47) | |
| 40 and below | 38 (81) |
| 41 - 49 | 3 (6) |
| 50 - 70 | 6 (13) |
| LACE[c] score (N=34) | |
| 0 - 4 | 2 (6) |
| 5 - 9 | 13 (38) |
| 10 and above | 19 (56) |

[a]NHYA: New York Heart Association.

[b]Although patients' initial evaluation was NYHA classification IV, this patient was included due to cardiologist recommendations.

[c]LACE: Length of stay, Acuity of admission, Comorbidity, and Emergency room usage.

### *Patients' Physical and Psychological Health Status and Health Care Behaviors*

Descriptive statistics for measures of physical and psychological health status and health care behaviors taken at the commencement of the program (baseline: within 0 - 14 d of enrollment) and after the intensive phase (follow-up: within 70 - 98 d of enrollment) are presented in Table 2. Thirty-one (77.5%) of the 40 patients with HF completed a first and follow-up KCCQ-12 assessment within the intensive phase of the VHHHFP. The mean summary score at follow-up was found to be significantly higher than the mean summary score at baseline. Consistent with this, the mean baseline and follow-up scores for the KCCCQ-12 subscale measures of physical limitations and quality of life were found to have significantly increased over time, as did the median subscale scores at baseline and follow-up for symptom frequency and social limitation.

**Table .** Descriptive statistics for patients' physical and psychological health status and health care behaviors at baseline and follow-up.

| Measure | Baseline period | Follow-up period | Statistic *t* (df) or *z* score | *P* value |
|---|---|---|---|---|
| KCCQ-12[a], N=31 | | | | |
| Summary score | 50.51 (SD 17.59) | 72.20 (SD 20.28) | *t*=–5.91 (30) | <.001 |
| Physical limitation | 47.09 (SD 29.77) | 69.43 (SD 22.62) | *t*=–4.02 (30) | <.001 |
| Symptom frequency | 60.40 (IQR 1-100) | 91.70 (IQR 35.40) | *z*=–4.43 | <.001 |
| Social limitation | 50.0 (IQR 1-100) | 82.50 (IQR 32.50) | *z*=–4.88 | <.001 |
| Quality of life | 43.75 (SD 21.71) | 62.91 (SD 25.73) | *t*=–3.94 (29) | <.001 |
| PHQ-4[b], N=32 | | | | |
| Total distress | 1.50 (IQR 0-7) | 0.0 (IQR 0-8) | *z*=–2.42 | <.02 |
| Anxiety | 1.0 (IQR 0-6) | 0.0 (IQR 0-4) | *z*=–2.53 | <.02 |
| Depression | 0.0 (IQR 0-4) | 0.0 (IQR 0-4) | *z*=–1.39 | =.16 |
| PAM-13[c], N=28 | 57.40 (SD 8.53) | 68.26 (SD 12.38) | *t*=–4.81 (27) | <.001 |

[a]KCCQ-12: Kansas City Cardiomyopathy Questionnaire.

[b]PHQ-4: Patient Health Questionnaire-4.

[c]PAM-13: Patient Activation Measure-13.

The PHQ-4 measure of psychological health was completed by 32 (80%) of program patients within the intensive phase of the VHHHFP. The median scores at baseline and follow-up for total distress (1.50, IQR 0-7 and 0.0, IQR 0-8; $z$=–2.42; $P$<.02), and the anxiety subscale (1.0, IQR 0-6 and 0.0, IQR 0-4; $z$=–2.53; $P$<.02) significantly reduced over time, while no statistically significant change over time was found in the median subscale score for depression (0.0, IQR 0-4 and 0.0, IQR 0-4; $z$=−1.39; $P$=.16). A total of 28 (70%) program patients completed the PAM-13. Analysis of the mean baseline (57.40 SD 8.53) and follow-up (68.26, SD 12.38) scores for the PAM-13 showed a statistically significant increase in patients' self-reported knowledge, beliefs, confidence, and skills about managing their HF throughout the intensive phase of the VHHHFP ($t_{27}$=−4.81; $P$<.001).

### *Patient Reported Experience Measures*

Of the 34 (85%) patients who responded to the PREMs questionnaire, all (100%) responded "always" or "mostly" to questions about their treatment and care (Table 3). Most patients (94%) responded that their views and concerns were always listened to. Twenty-six (76%) patients responded that they always knew how to recognize HF or heart attack symptoms and what to do next. Seven (21%) patients responded that they mostly knew how to recognize HF, and one responded that they only sometimes knew how to recognize HF symptoms. In relation to the patients' experience of the technology of the VHHHFP, most patients were satisfied to varying degrees. Similarly, most patients (n=33, 97%) were satisfied to very satisfied that they had the knowledge required to use the technology.

**Table .** Frequency of patient-reported experience measures responses.

| Scale and experience item | Response 1, n (%) | Response 2, n (%) | Response 3, n (%) | Response 4, n (%) | Response 5, n (%) | Total, n (%) |
|---|---|---|---|---|---|---|
| Response: 1=always, 2=mostly, 3=sometimes, 4=rarely, and 5=never | | | | | | |
| My views and concerns were listened to. | 32 (94) | 0 (0) | 1(3) | 0 (0) | 1 (3) | 34 (100) |
| My individual needs were met. | 30 (88) | 0 (0) | 4(12) | 0 (0) | 0 (0) | 34 (100) |
| I felt cared for. | 31 (91) | 3 (9) | 0 (0) | 0 (0) | 0 (0) | 34 (100) |
| I was involved as much as I wanted in making decisions about my treatment and care. | 30 (88) | 4 (12) | 0 (0) | 0 (0) | 0 (0) | 34 (100) |
| I was kept informed as much as I wanted about my treatment and care. | 33 (97) | 1 (3) | 0 (0) | 0 (0) | 0 (0) | 34 (100) |
| The staff involved in my care communicated with each other about my treatment. | 31 (91) | 3 (9) | 0 (0) | 0 (0) | 0 (0) | 34 (100) |
| I knew how to recognize heart failure or heart attack symptoms, and what to do next if I experienced symptoms. | 26 (76) | 7 (21) | 1 (3) | 0 (0) | 0 (0) | 34 (100) |
| I felt confident in the safety of my treatment and care. | 30 (88) | 4 (12) | 0 (0) | 0 (0) | 0 (0) | 34 (100) |
| Response: 1=very satisfied, 2=satisfied, 3=neither, 4=dissatisfied, and 5=very dissatisfied | | | | | | |
| How satisfied are you that the technology operated as expected? (ie, tablet, app, biometric devices, and video calls) | 22 (65) | 10 (30) | 2 (6) | 0 (0) | 0 (0) | 34 (100) |
| How satisfied are you that you have the knowledge to use the technology? (ie, tablet, app, biometric devices, and video calls) | 26 (76) | 7 (21) | 1 (3) | 0 (0) | 0 (0) | 34 (100) |

### *Readmission to Hospital*

The analysis of the patients' admission to hospital during the intensive phase of the VHHHFP was undertaken on the 40 patients who completed the intensive phase of the VHHHFP within 70 - 98 days and those patients who withdrew from the program (n=7) or failed to complete the intensive phase of the program within 70 - 98 days (n=2). A total of 6 hospital admissions were recorded for 5 patient participants (10.2% of 49 program participants) within 30 days of the patients' commencement of the VHHHFP (based on the patient's recorded start or enrollment date). One patient was admitted twice in 3 days within the 30 days. The earliest admission within the 30 days occurred at 7.8 days, the latest at 27.9 days.

A further 8 hospital admissions for 7 patients (14.2% of participants) were recorded between days 31 and 98 of the patients' commencement of the VHHHFP (Table 4). None of these patients had previously recorded an admission within 30 days. One patient was admitted twice in 3 days within the 31 -

to 98-day period. The earliest admission during this period occurred at 36.7 days, the latest at 89.9 days. In total, 14 separate hospital admissions were recorded during the intensive phase of the VHHHFP across 12 patients (24.5% of program participants). However, the data provided were not complete and consistent for all hospital admissions to determine the reason or cause for the admission (ie, all-cause, HR-related, or otherwise) for the 14 hospital admissions.

**Table .** Frequency of all causes and 30-day readmissions to the hospital during the intensive phase of the Virtual Home Health Heart Failure Program.

| Frequency of hospital admissions | Number of patients, n (%) | Readmitted within 30 days of commencing the VHHHFP[a], n (%) |
|---|---|---|
| No admissions | 37 (75.5) | 44 (89.8) |
| 1 | 10 (20.4) | 4 (8.2) |
| 2 | 2 (4.1) | 1 (2.0) |

[a]VHHHFP: Virtual Home Health Heart Failure Program.

### Semistructured Interviews

A total of 9 patients participated in the interviews, 4 females and 5 males, ranging in age from 53 to 82 years. The VHHHFP was described by patients as a virtual resource that facilitated clinician interaction and care and was a platform for sharing clinical patient self-measures and communication. Three themes were identified from the analysis that described the patient experiences of the VHHHFP following completion of the 12-week intensive phase of the VHHHFP (Table S1 in Multimedia Appendix 2). The first theme described how the VHHHFP enhanced patient capability for self-management of HF. Patients described recognizing the need to engage in individualized self-care, and the program enhanced their ability and empowerment to confidently manage their HF. The second theme, improved patient comfort, was an outcome of engagement with the program, where patients described the VHHHFP as allaying patient fear and uncertainty regarding their HF condition, and the information and education (provided from the program) contributed to patient comfort and support from family. The third theme described positive influences on calm and how calm improved through coordinated care and a supportive environment. The virtual program contributed to this supportive environment.

A total of 6 health care professionals (4 nurses, 1 cardiologist, and 1 dietitian) shared their experiences and perceptions of the VHHHFP. For the study, the term “staff” will be used; however, in the VHHHFP, staff were either used directly by the organization delivering the service or as external providers. Through individual interviews, the experiences and perceptions of HCPs involved in the VHHHFP were explored (Table S2 in Multimedia Appendix 1). The first theme identified was health care professionals’ improving patient capability through a shared understanding of health needs. This theme included creating a supportive environment of care, the importance of guidelines for shared care, and health care professionals getting satisfaction from supporting patients in a virtual model of care. The second theme, improving capability through care practices, encompassed staff perception of making a difference to patient self-care as an outcome of the VHHHFP. This was described as achieved through the provision of care to maximize outcomes and patient capability and empowering patients in self-care practices. The third theme, promoting comfort and calm through virtual coordinated and collaborative care approach, described how staff identified that the VHHHFP contributed to patient comfort and calm. Recognition of the multidisciplinary model of care and that the virtual program enables partnership with the care team and patients was identified as critical components to the success of the program. Some participants also acknowledged that access to GPs or cardiologists presented a challenge sometimes. Table S2 in Multimedia Appendix 2 provides exemplars for the themes and subthemes.

## Discussion

### Principal Findings

This study contributes to the knowledge base on the impact of a virtual health approach from both a patient and clinician perspective and on how virtual health solutions can be integrated into existing care. A statistically significant improvement in physical health and well-being on completion of the intensive phase was noted as measured by the KCCQ-12. Drawing on previous research as a reference point [24], a 5-point threshold for meaningful clinical change has been proposed as equivalent to a ~10% relative reduction in the risk of adverse clinical events. Using this interpretation, a mean improvement of 21.7 in KCCCQ-12 patient summary scores from commencement to completion of the intensive phase of the VHHHFP could be considered to represent a 40% relative reduction in the risk of adverse clinical events pre- and post-intensive phase of the VHHHFP [24]. This highlights the beneficial effect of the program on patients’ physical health and well-being. The positive impact of postdischarge interventions on KCCQ-12 scores and, therefore, the well-being of patients with HF has also been demonstrated elsewhere. An intervention study by Stubblefield et al [31] comprising self-care activities, home visits, and telephone calls to coach participants in the aspects of HF self-management demonstrated a 5.4-point increase in KCCCQ-12 scores in the intervention group compared to the usual structured care group. The ongoing connection with clinicians following discharge from the hospital appears to contribute to the well-being of patients with HF. The quantitative findings reflect the qualitative themes of patients feeling “empowered to manage self-care activities,” that “information and education contributed to patient comfort,” and that the “virtual program provided a supportive environment.”

## Comparison to Prior Work

Anxiety and depression are known to be significant issues in patients with HF [30], impacting many areas of a patient's life, including adherence to treatment plans. Anxiety and depression are also associated with reduced quality of life [32], reduced exercise capacity [33], and increased hospitalizations and mortality [34]. In this study, there was a statistically significant reduction in self-reported levels of anxiety and improvement in self-reported symptoms on the distress scale during the program, highlighting the importance of routinely monitoring mood in patients with HF. Enabling patients to self-report symptoms of anxiety and depression safely provides the opportunity for early recognition and management, which in turn may improve the management of HF and improve patient outcomes. This was reinforced in the qualitative data of this study, where patients reported feeling reassured when they had a clinician to contact if and when they had an exacerbation of their symptoms. Studies of patients with cardiac conditions have traditionally measured easily accessible outcomes such as hospitalizations and mortality. A strength of this study is the focus on patient-reported outcomes, which provides clinicians and researchers with an accurate report of health status directly from the patient, leading to the capture of meaningful data on the patient experience [35].

Enhanced patient capability is an important outcome for the program given the centrality of self-management in HF. This was highlighted in the qualitative results relating to patient experiences after engaging with the VHHHFP. The availability of staff to ask questions and allay concerns was described by patients as important and also contributing to their comfort, providing reassurance that someone was monitoring their measurements. This highlights the value of the availability of support beyond scheduled consultations. The CCC framework highlights the component of calm, which, from a patient's perspective, reiterates the important benefits of having services designed around them rather than around the HCP, reducing the stress of accessing care and minimizing disruption to their lives. These qualitative findings could potentially be linked to the positive change in the PHQ-4 total distress score, which was statistically significant. Virtual programs such as the VHHHFP promote outcomes related to calm for postacute support services. The impact and importance of care integration have been evidenced in other studies and populations where continuity and availability of health care providers were recommended [36]. The findings from this study highlighted the need to ensure all HCPs involved in the patient's care were familiar with the program's services to ensure smooth communication processes. The impact of communication processes on calm has also been reported in other populations. In a study with people diagnosed with bipolar disorder, participants reported that their sense of calm was enhanced by increased engagement time and improved communication with health care providers [37], and in young adults diagnosed with cancer [38].

The qualitative data support a clinically significant improvement in self-reported knowledge, beliefs, confidence, and skills about managing HF throughout the intensive phase of the VHHHFP. Similar effects were reported in a study of patients with atrial fibrillation undertaking a virtual program during the COVID-19 pandemic. Improvements in self-monitoring abilities and self-management behaviors, and statistically significant reductions in anxiety and depression were also findings from this study [39]. Patients with HF are particularly vulnerable in the immediate post hospital discharge period while transitioning to their home environments.

The up-titration of guideline-directed medications is a cornerstone in patients with HF, particularly in the context of reduced ejection fraction [40], and is an essential strategy to fill the gap in care during the early discharge period [41]. Evidence-based clinical guidelines recommend that each medication be titrated as tolerated to the target dose, which was supported by landmark clinical trials to achieve maximum benefits [42]. Medication up-titration had limited uptake in this study despite processes for this being in place, including the availability of medication up-titration request forms. The qualitative data did not directly identify why the opportunity for up-titration was not readily taken up, but this could be related to the references to challenges in accessing GPs or cardiologists, where any changes to a titration plan required a GP or specialist oversight. The qualitative data indicate that nursing staff reported the need to titrate medications but did not have the protocol or scope to do so without contacting the patients' GP or specialist. In a study exploring the barriers to up-titration of beta-blockers in patients with HF in the community, barriers identified included physicians' concerns about medication side effects and polypharmacy, existing health care system barriers, comorbidities, patient communication, and physicians' knowledge and experience [43]. The lack of uptake and potential missed opportunity to improve patient outcomes is an area for further exploration. An extended scope of practice for nurses to up-titrate and initiate guideline-directed medications could be an option for future iterations of the VHHHFP [44].

Data on hospital admissions were not collected from the time of discharge from the hospital, but from the time of enrollment into the program, which is a limitation of this study. However, the rate of hospital admissions within 30 days of enrollment into the VHHHFP was 10.2%, which does compare favorably to 30-day all-cause readmission rates in a previous study, which demonstrated 20% [45]. While a direct comparison cannot be made due to potential delays from discharge from the hospital to onboarding to the program, the results are still promising in that readmission was well below 20%. The finding is also in line with previous research that demonstrated a reduction in readmission rates in similar, virtually delivered, remote monitoring programs for patients with HF [14].

## Strengths and Limitations

A strength of this study was the exploration of self-reported patient outcomes and patient and clinical staff perceptions of the VHHHFP. The exploratory qualitative process evaluation provided valuable insights into the acceptability and usability of the intervention from the perspectives of the participants. The qualitative data allowed a deeper understanding of how participants responded to the program and the contextual factors that influenced the study outcomes. Trials in the study of patients with cardiac care routinely report outcomes such as

hospitalizations and mortality. However, a significant strength of this study was the focus on patient-reported outcomes, an approach that could be more widely adopted in virtual health programs and beyond.

It is important to acknowledge the limitations that may affect the validity and generalizability of the research findings. First, the service was set up as a pilot to assess feasibility, and as such, the sample size was not powered to detect change with a level of statistical certainty. Second, the evaluation commenced after the commencement of the service, and as such, it was not possible to create a control group, minimizing the opportunity to assess selection or detection bias. The variation in the time taken to complete the intensive phase and the submission of data associated with key time points made comparisons between participants in the current evaluation challenging. To include as many participants as possible in the study, the parameters relating to the time taken to complete the intensive phase and the submission of data associated with a given time point were extended. Third, the number of patient participants who agreed to be interviewed postcompletion of the 12-week intensive program was low, and there may have been a self-selection bias among patients who chose to participate in the qualitative component of this evaluation. In addition, direct comparisons with data on readmission rates could not be made where the date of discharge from the hospital (the admission that led to the initial referral to the program) was not collected.

### Future Directions

The growth in virtual health programs, such as VHHHFP, has demonstrated a range of benefits to patients regarding improved access to advice and guidance on their medical condition without the need to visit health care facilities. This approach has led to the reallocation of much-needed health care resource provision [46], especially within the high-resource area of HF management. The virtual nature of the intervention creates the opportunity to scale within and across health care services. However, this service sits in what can be termed the “missing middle” of health care, with the service spanning a gap between care provided in the acute care and community settings. One of the key barriers to further implementation in the Australian context relates to funding mechanisms. Ongoing support for services that do not meet existing health care funding mechanisms can be uncertain, limiting their increased uptake and opportunity for further translation of benefits to a broader section of the community. Further research into enhancing the adoption of such models is needed. The use of technologies to support the management of HF is growing internationally. Understanding how virtual health care that uses remote patient monitoring can be integrated into existing systems and models of care is a challenge that requires multilevel collaboration. The findings of this study support the need to develop and adopt virtual health care solutions for chronic disease management, including and beyond HF.

## Conclusion

The evaluation of the VHHHFP demonstrates improvements in both clinical and behavioral outcomes, directly addressing our primary study aim. Patients completing the program showed statistically significant improvements in all KCCQ-12 domains, including physical limitations, quality of life, symptom frequency, and social limitations. Psychological health measures similarly improved with reductions in total distress and anxiety scores. Our second aim, exploring participant experiences, revealed enhanced capability, improved comfort, and positive influences on calm according to the VBHC framework. Clinician experiences identified benefits in patient capability through shared decision-making and care practices, while also noting virtual coordination promoted patient comfort and calm. This evaluation, using both quantitative and qualitative methods, provides evidence for the effectiveness of the VHHHFP model in HF management.

## Acknowledgments

We would like to extend our thanks to the patients and staff for their participation in this research. This work was supported with funding from Ramsay Connect, Ramsay Health Care Australia.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

NM, BE, RS, LW, KS, JY, and HM contributed to the conceptualization of the study. Data curation was performed by HM, SR, and PP. SR, PP, HM, NM, and RS conducted the formal analysis. Funding acquisition was led by BE, NM, LW, and RS. HM and SR were responsible for the investigation. BE, NM, LW, RS, KS, and JY developed the methodology. BE and NM oversaw project administration. BE, NM, LW, and RS provided resources. Supervision was carried out by BE, LW, RS, and NM. PP and SR performed validation. Visualization was completed by NM, MR, and PP. NM, BE, LW, KS, RS, and PP drafted the original manuscript.

## Conflicts of Interest

JY and MR are employees of Home Health and Australian Unity, and KS is an employee of Ramsay Health Care (Ramsay Connect), which are the companies involved in the development and delivery of VHHHFP. The authors have no other conflicts of interest to declare.

Multimedia Appendix 1
Open-ended participant interview questions.
[DOCX File, 25 KB - cardio_v9i1e64877_app1.docx ]

Multimedia Appendix 2
Themes and exemplar quotes generated from patient and health care professional interviews.
[DOCX File, 50 KB - cardio_v9i1e64877_app2.docx ]

## References


1. Luiso D, Herrero-Torrus M, Badosa N, et al. Quality of life in older patients after a heart failure hospitalization: results from the SENECOR study. J Clin Med 2022 May 27;11(11):3035. [doi: 10.3390/jcm11113035] [Medline: 35683423]
2. Savarese G, Becher PM, Lund LH, et al. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res 2023 Jan 18;118(17):3272-3287. [doi: 10.1093/cvr/cvac013]
3. Becher PM, Lund LH, Coats AJS, Savarese G. An update on global epidemiology in heart failure. Eur Heart J 2022 Aug 21;43(32):3005-3007. [doi: 10.1093/eurheartj/ehac248] [Medline: 35578978]
4. Heidenreich PA, Fonarow GC, Opsha Y, et al. Economic issues in heart failure in the United States. J Card Fail 2022 Mar;28(3):453-466. [doi: 10.1016/j.cardfail.2021.12.017] [Medline: 35085762]
5. Shahim B, Kapelios CJ, Savarese G, Lund LH. Global public health burden of heart failure: an updated review. Card Fail Rev 2023;9:e11. [doi: 10.15420/cfr.2023.05] [Medline: 37547123]
6. Li Y, Cao GY, Jing WZ, Liu J, Liu M. Global trends and regional differences in incidence and mortality of cardiovascular disease, 1990−2019: findings from 2019 global burden of disease study. Eur J Prev Cardiol 2023 Feb 14;30(3):276-286. [doi: 10.1093/eurjpc/zwac285] [Medline: 36458973]
7. Bozkurt B, Ahmad T, Alexander KM, et al. Heart failure epidemiology and outcomes statistics: a report of the Heart Failure Society of America. J Card Fail 2023 Oct;29(10):1412-1451. [doi: 10.1016/j.cardfail.2023.07.006] [Medline: 37797885]
8. Khan MS, Shahid I, Bennis A, et al. Global epidemiology of heart failure. Nat Rev Cardiol 2024 Oct;21(10):717-734. [doi: 10.1038/s41569-024-01046-6] [Medline: 38926611]
9. Heart, stroke and vascular disease: Australian facts. Australian Institute of Health and Welfare. 2023. URL: https://www.aihw.gov.au/reports/heart-stroke-vascular-diseases/hsvd-facts/report-editions [accessed 2025-06-11]
10. Dang T, Chan W, Khawaja S, et al. Hospital costs for unplanned re-admissions within 30 days of hospitalisations with heart failure, Australia, 2013-2017: a retrospective cohort study. Med J Aust 2024 Sep 16;221(6):317-323. [doi: 10.5694/mja2.52424] [Medline: 39188208]
11. Parsons RW, Liew D, Neville AM, et al. The epidemiology of heart failure in the general Australian community - study of heart failure in the Australian primary carE setting (SHAPE): methods. BMC Public Health 2020 May 11;20(1):648. [doi: 10.1186/s12889-020-08781-8] [Medline: 32393222]
12. Wong-Rieger D. Values-based and outcomes-based healthcare delivery. Australian Pharmacist 2014;33(3):20. [doi: 10.1371/journal.pone.0258674] [Medline: 34648595]
13. Cahn A, Akirov A, Raz I. Digital health technology and diabetes management. J Diabetes 2018 Jan;10(1):10-17. [doi: 10.1111/1753-0407.12606] [Medline: 28872765]
14. Clemente MRC, Felix N, Navalha DDP, et al. Long-term impact of home-based monitoring after an admission for acute decompensated heart failure: a systematic review and meta-analysis of randomised controlled trials. EClinicalMedicine 2024 May;71:102541. [doi: 10.1016/j.eclinm.2024.102541] [Medline: 38545427]
15. Zheng Y, Adhikari S, Li X, et al. Association between video-based telemedicine visits and medication adherence among patients with heart failure: retrospective cross-sectional study. JMIR Cardio 2024 Dec 5;8:e56763. [doi: 10.2196/56763] [Medline: 39637412]
16. Bente BE, Wentzel J, Schepers C, et al. Implementation and user evaluation of an eHealth technology platform supporting patients with cardiovascular disease in managing their health after a cardiac event: mixed methods study. JMIR Cardio 2023 Mar 24;7:e43781. [doi: 10.2196/43781] [Medline: 36961491]
17. Lee KCS, Breznen B, Ukhova A, et al. Virtual healthcare solutions in heart failure: a literature review. Front Cardiovasc Med 2023;10:1231000. [doi: 10.3389/fcvm.2023.1231000] [Medline: 37745104]
18. Wali S, Guessi Margarido M, Shah A, et al. Expanding telemonitoring in a virtual world: a case study of the expansion of a heart failure telemonitoring program during the COVID-19 pandemic. J Med Internet Res 2021;23(1):e26165. [doi: 10.2196/26165] [Medline: 33444153]
19. Friedman DM, Goldberg JM, Molinsky RL, et al. A virtual cardiovascular care program for prevention of heart failure readmissions in a skilled nursing facility population: retrospective analysis. JMIR Cardio 2021;5(1):e29101. [doi: 10.2196/29101] [Medline: 34061037]
20. Creswell JW. A Concise Introduction to Mixed Methods Research: Sage Publications; 2021. URL: http://au.sagepub.com/en-gb/oce/a-concise-introduction-to-mixed-methods-research/book266037

21. Barnow BS, Pandey SK, Luo QE. How mixed-methods research can improve the policy relevance of impact evaluations. Eval Rev 2024 Jun;48(3):495-514. [doi: 10.1177/0193841X241227480] [Medline: 38299483]
22. Lee SYD, Iott B, Banaszak-Holl J, et al. Application of mixed methods in health services management research: a systematic review. Med Care Res Rev 2022 Jun;79(3):331-344. [doi: 10.1177/10775587211030393] [Medline: 34253078]
23. Burns DJP, Arora J, Okunade O, et al. International Consortium for Health Outcomes Measurement (ICHOM): standardized patient-centered outcomes measurement set for heart failure patients. JACC Heart Fail 2020 Mar;8(3):212-222. [doi: 10.1016/j.jchf.2019.09.007] [Medline: 31838032]
24. Spertus JA, Jones PG, Sandhu AT, Arnold SV. Interpreting the Kansas City Cardiomyopathy Questionnaire in clinical trials and clinical care: JACC State-of-the-Art Review. J Am Coll Cardiol 2020 Nov 17;76(20):2379-2390. [doi: 10.1016/j.jacc.2020.09.542] [Medline: 33183512]
25. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
26. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026. [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
27. Patient experience. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/our-work/indicators-measurement-and-reporting/australian-hospital-patient-experience-question-set [accessed 2024-07-28]
28. Rajaguru V, Han W, Kim TH, et al. LACE index to predict the high risk of 30-day readmission: a systematic review and meta-analysis. J Pers Med 2022 Mar 30;12(4):545. [doi: 10.3390/jpm12040545] [Medline: 35455661]
29. Teisberg E, Wallace S, O'Hara S. Defining and implementing value-based health care: a strategic framework. Acad Med 2020 May;95(5):682-685. [doi: 10.1097/ACM.0000000000003122] [Medline: 31833857]
30. Rashid S, Qureshi AG, Noor TA, et al. Anxiety and depression in heart failure: an updated review. Curr Probl Cardiol 2023 Nov;48(11):101987. [doi: 10.1016/j.cpcardiol.2023.101987] [Medline: 37473952]
31. Stubblefield WB, Jenkins CA, Liu D, et al. Improvement in Kansas City Cardiomyopathy Questionnaire scores after a self-care intervention in patients with acute heart failure discharged from the emergency department. Circ Cardiovasc Qual Outcomes 2021 Oct;14(10):e007956. [doi: 10.1161/CIRCOUTCOMES.121.007956] [Medline: 34555929]
32. AbuRuz ME. Anxiety and depression predicted quality of life among patients with heart failure. J Multidiscip Healthc 2018;11(null):367-373. [doi: 10.2147/JMDH.S170327] [Medline: 30104881]
33. Chialà O, Vellone E, Klompstra L, et al. Relationships between exercise capacity and anxiety, depression, and cognition in patients with heart failure. Heart Lung 2018 Sep;47(5):465-470. [doi: 10.1016/j.hrtlng.2018.07.010]
34. Celano CM, Villegas AC, Albanese AM, et al. Depression and anxiety in heart failure: a review. Harv Rev Psychiatry 2018;26(4):175-184. [doi: 10.1097/HRP.0000000000000162] [Medline: 29975336]
35. Thomson RJ, Warren A, Pimento S, et al. In perspective: the patient at the heart of research in acute cardiovascular care. Eur Heart J Acute Cardiovasc Care 2023 May 4;12(5):340-343. [doi: 10.1093/ehjacc/zuad027] [Medline: 36938596]
36. Christian N, McFall C, Suarez J, et al. Achieving calm: a study on the health care experiences of people with lived experience of homelessness in central Texas. Qual Health Res 2022 Dec;32(14):2118-2125. [doi: 10.1177/10497323221135795] [Medline: 36356263]
37. Siegel-Ramsay JE, Sharp SJ, Ulack CJ, et al. Experiences that matter in bipolar disorder: a qualitative study using the capability, comfort and calm framework. Int J Bipolar Disord 2023 Apr 20;11(1):13. [doi: 10.1186/s40345-023-00293-9] [Medline: 37079153]
38. Guzik P, McKinney JA, Ulack C, et al. Outcomes that matter most to young adults diagnosed with Cancer: a qualitative study. J Adolesc Young Adult Oncol 2021 Oct;10(5):534-539. [doi: 10.1089/jayao.2020.0150] [Medline: 33237848]
39. Rosman L, Armbruster T, Kyazimzade S, et al. Effect of a virtual self-management intervention for atrial fibrillation during the outbreak of COVID-19. Pacing Clin Electrophysiol 2021 Mar;44(3):451-461. [doi: 10.1111/pace.14188] [Medline: 33565642]
40. Williams R, Donald R, James A, et al. Evaluation of a prescribing pharmacist-led heart failure (HF) up-titration clinic. Eur Heart J 2022 Oct 3;43(Supplement_2). [doi: 10.1093/eurheartj/ehac544.1035]
41. Cox ZL, Nandkeolyar S, Johnson AJ, et al. In-hospital initiation and up-titration of guideline-directed medical therapies for heart failure with reduced ejection fraction. Card Fail Rev 2022 Jan;8:e21. [doi: 10.15420/cfr.2022.08] [Medline: 35815257]
42. Greene SJ, Fonarow GC, DeVore AD, et al. Titration of medical therapy for heart failure with reduced ejection fraction. J Am Coll Cardiol 2019 May;73(19):2365-2383. [doi: 10.1016/j.jacc.2019.02.015] [Medline: 30844480]
43. Levitan EB, Van Dyke MK, Loop MS, et al. Barriers to beta-blocker use and up-titration among patients with heart failure with reduced ejection fraction. Cardiovasc Drugs Ther 2017 Dec;31(5-6):559-564. [doi: 10.1007/s10557-017-6764-8] [Medline: 29181610]
44. Güder G, Störk S, Gelbrich G, et al. Nurse-coordinated collaborative disease management improves the quality of guideline-recommended heart failure therapy, patient-reported outcomes, and left ventricular remodelling. Eur J Heart Fail 2015 Apr;17(4):442-452. [doi: 10.1002/ejhf.252] [Medline: 25727879]

45. Al-Omary MS, Davies AJ, Evans TJ, et al. Mortality and readmission following hospitalisation for heart failure in Australia: a systematic review and meta-analysis. Heart Lung Circ 2018 Aug;27(8):917-927. [doi: 10.1016/j.hlc.2018.01.009] [Medline: 29519691]
46. Gorodeski EZ, Goyal P, Cox ZL, et al. Virtual visits for care of patients with heart failure in the era of COVID-19: a statement from the Heart Failure Society of America. J Card Fail 2020 Jun;26(6):448-456. [doi: 10.1016/j.cardfail.2020.04.008] [Medline: 32315732]

## Abbreviations

**CCC:** Capability, Comfort, and Calm
**CNS:** clinical nurse specialist
**GP:** general practitioner
**HCP:** health care practitioner
**HF:** heart failure
**KCCQ-12:** Kansas City Cardiomyopathy Questionnaire-12
**LACE:** Length of stay, Acuity of admission, Comorbidity, and Emergency room usage
**PAM-13:** Patient Activation Measure-13
**PHQ-4:** Patient Health Questionnaire-4
**PREM:** Patient Reported Experience Measure
**VBHC:** value-based health care
**VHHHFP:** Virtual Home Health Heart Failure Program

*Edited by A Coristine; submitted 30.07.24; peer-reviewed by A Adekoya, A Hidki, M Gasmi; revised version received 24.04.25; accepted 24.04.25; published 23.07.25.*

*Please cite as:*
*McKay N, Saunders R, Metcalfe H, Robinson S, Palamara P, Steer K, Yoo J, Ranogajec M, Whitehead L, Ewens B*
*Evaluation of a Virtual Home Health Heart Failure Program: Mixed Methods Study*
*JMIR Cardio 2025;9:e64877*
*URL: https://cardio.jmir.org/2025/1/e64877*
*doi:10.2196/64877*

# Barriers and Enablers to Routine Clinical Implementation of Cardiac Implantable Electronic Device Remote Monitoring in Australia Among Cardiologists, Cardiac Physiologists, Nurses, and Patients: Interview Study

Brodie Sheahen[1], MD; Edel T O'Hagan[1], PhD; Kenneth Cho[1], MBBS; Tim Shaw[1,2], PhD; Astin Lee[3,4], MBBS; Sean Lal[2,5], MBBS; Aaron L Sverdlov[6], MBBS; Clara Chow[1,7], MBBS

[1]Westmead Applied Research Center, The University of Sydney, Westmead, Australia

[2]Charles Perkins Centre, The University of Sydney, Camperdown, Australia

[3]Cardiology Department, The Wollongong Hospital, Wollongong, Australia

[4]School of Medicine, Faculty of Science Medicine and Health, University of Wollongong, Wollongong, Australia

[5]Cardiology Department, Royal Prince Alfred Hospital, Sydney, Australia

[6]Cardiology Department, John Hunter Hospital, Newcastle, Australia

[7]Cardiology Department, Westmead Hospital, Westmead, Australia

**Corresponding Author:**
Brodie Sheahen, MD
Westmead Applied Research Center, The University of Sydney, Westmead, Australia

## Abstract

**Background:** Remote monitoring (RM) of cardiac implantable electronic devices (CIEDs) has demonstrated many patient and health care system benefits. Consequently, the use of RM technology for patients with CIEDs is the standard of care as highlighted by international guidelines. However, RM has not yet been integrated into universal, routine clinical practice.

**Objective:** We aimed to establish key stakeholder perspectives on the barriers and enablers of CIED RM implementation and to apply the theoretical domain framework to highlight the most effective approaches to facilitate routine adoption of CIED RM.

**Methods:** This was a qualitative study, using semistructured interviews to explore the barriers and enablers encountered when incorporating RM into CIED management. Participants included cardiologists, cardiac clinicians or physiologists, nurses, and patients. Interviews were transcribed verbatim and analyzed through inductive thematic analysis and deductive approaches using the NVivo (version 14; QRS International Pty Ltd) software. The theoretical domains framework was used to understand barriers and enablers. In the inductive phase, we did not assess trustworthiness, as our thematic analysis approach views data as interpretations rather than objective truths. In the deductive phase, we conferred to ensure consistency in theme alignment with existing frameworks.

**Results:** Interviews were conducted among 35 participants (16 patients, 10 cardiologists, and 9 cardiac physiologists and nurses). We identified 5 main themes and their associated subthemes, with 1 representing an enabler and 4 representing barriers. They were: (1) patient benefits from RM, such as improved CIED and cardiovascular management, and improved patient-centered care; (2) insufficient allocation of CIED RM resources, which included insufficient RM clinic funding and staffing, insufficient RM service reimbursement, and RM infrastructure and access inequity; (3) suboptimal management of data, which includes inconsistent RM alert interpretation and management, lack of guidance for clinic staff on RM data management, and an increased alert burden for clinics; (4) insufficient patient education post-CIED implant, this was attributed to limited health care worker availability and resulted in inadequate patient CIED and RM knowledge postimplant and patient anxiety associated with RM; and (5) patient engagement with CIED management, which included the need for increased patient interaction with RM alerts and the ability to share data with patients. These subthemes were mapped to 6 specific domains of the theoretical domains framework: “Beliefs About Capabilities,” “Environmental Context and Resources,” “Beliefs About Consequences,” “Knowledge,” “Emotions,” and “Goals.”

**Conclusions:** Patient engagement was identified in 3 of the 5 themes describing barriers and enablers to RM. These highlight the importance of addressing patient engagement with RM to better implement and integrate the use of RM into routine clinical practice. Barriers and enablers extend across multiple domains and suggest that a multipronged approach is required to translate the gold standard care of RM to routine clinical practice.

***(JMIR Cardio 2025;9:e67758)*** doi:10.2196/67758

**KEYWORDS**

remote monitoring; cardiac implantable electronic devices; cardiac implant; patient engagement; barriers and enablers; cardiovascular disease; CVD; congestive heart failure; CHF; myocardial infarction; MI; unstable angina; angina; cardiac arrest; atherosclerosis; cardiology; cardiologist

## Introduction

The use of remote monitoring (RM) is the standard of care for patients with cardiac implantable electronic devices (CIEDs) and is poised for wider adoption in the coming years, backed by growing endorsements from large cardiac societies such as the Heart Rhythm Society and Cardiac Society of Australia and New Zealand [1,2]. Whilst this uptake in RM is a positive move for improving patient care, in turn, it raises concerns about the capacity of device clinics to manage the associated workload [3,4]. Recent studies have estimated that managing 1000 patients with CIED with RM necessitates a workforce commitment of approximately 30 - 46 hours per week by the clinical team [5].

The relative novelty of the technology creates challenges when incorporating CIED RM into clinical practice. Insufficient funding, lack of appropriate infrastructure, and lack of standardized workflow are commonly cited barriers [3,4,6,7]. Furthermore, despite some cardiac organizations placing a greater emphasis on patient engagement in the CIED, engagement initiatives are lacking, particularly surrounding patient education and information delivery [8,9]. The research to date suggests that implementation of RM requires cohesive management among many stakeholders, such as cardiologists, nurses, cardiac physiologists, and patients.

It is recognized across multiple sectors of health care that effective and sustainable implementation of research and innovations into clinical care relies on relevant stakeholders' input into the integration of the intervention [10]. A comprehensive implementation analysis of RM across all relevant stakeholders has not been conducted internationally. Currently, there is a scarcity of information on stakeholder perspectives of the barriers and enablers of CIED RM. Thus, this study aimed to (1) establish broad stakeholder perspectives on issues surrounding the routine implementation of CIED RM and (2) apply the theoretical domain framework to highlight, through an implantation science lens, the most effective approaches to facilitate routine adoption of CIED RM.

## Methods

### Study Overview

This was a qualitative study, using semistructured interviews to explore individual perspectives on barriers and facilitators to RM of CIEDs and patient engagement. This study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) [11] checklist for study execution and subsequent reporting.

### Theoretical Domains Framework

We used the theoretical domains framework (TDF) to understand barriers and enablers through an implementation science lens. The TDF is comprised of 14 domains and 84 constructs to bring together many behavior-change theories. It was designed to bridge the gap between behavior-change theory and various medical disciplines, making it both accessible and applicable to a wide range of health care professionals [12].

### Research Team and Reflexivity

We adopted a hybrid approach, combining postpositivist principles and codebook thematic analysis [13]. This approach recognizes that knowledge is never fully objective but integrates procedures to ensure rigor. Consistent with this perspective, we acknowledge that all observations are shaped by the researcher's perspectives, assumptions, and contexts, which are tentative and subject to revision. The research team was composed of cardiologists (CC, AL, SL, AS, and KC), a doctor-in-training and PhD student (BS), clinical researchers (ETO, CC, AL, SL, AS, and KC), and a digital health expert (TS). Researchers (BS, EO, CC, and TS) have experience conducting qualitative research, while clinician-researchers (CC, AL, SL, AS, and KC) have clinical cardiology experience. Interviews were conducted by the lead researcher (BS). Participants were aware that the interviewer was a PhD student and doctor-in-training; however, they had not met him prior to their interview.

### Study Setting and Recruitment

Between July 2022 and April 2023, we identified stakeholders (cardiologists, cardiac physiologists, nurses, and patients) who either used RM or were involved in analyzing and deciding on appropriate action for the data or alerts received via CIED RM. All stakeholders were based in Australia. Australia's health care system combines Medicare**,** which provides universal public coverage, in parallel with private insurance for additional services. Stakeholders were recruited from 5 hospitals providing CIED and at least some RM services to urban and regional areas of New South Wales, Australia: Westmead, Wollongong, Royal Prince Alfred, Concord, and John Hunter. Stakeholder eligibility criteria included being 18 years or older and English speaking. Patient-specific criteria included currently having a CIED in-situ, which is undergoing RM. Cardiologist-specific criteria included being a consultant, public hospital or private practice-based, and managing at least one patient currently receiving RM. Cardiac physiologist and nurse-specific criteria included managing at least one patient currently receiving RM and public hospital or private practice-based.

### Procedure or Data Collection

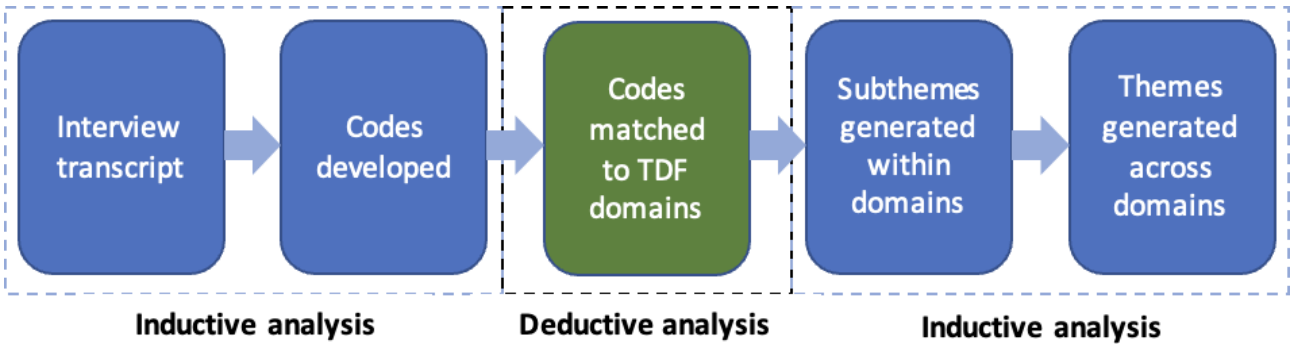
Specific interview guides (Multimedia Appendix 1) were developed based on the stakeholders being interviewed (cardiologists, cardiac physiologists, nurses or allied health clinicians, and patients). The interview guides explored (1) stakeholder perspectives on the barriers and facilitators of CIED RM and (2) patient engagement with CIED and overall cardiovascular disease (CVD) management. Additionally, participant demographic data were collected verbally at the beginning of each interview. To develop the interview guides, we conducted a comprehensive literature review, identifying

relevant studies on patient perspectives and existing interview guides used in similar studies. Interview guides were further refined after consulting with a cardiologist and conducting pilot interviews to ensure that questions were clear, comprehensive, and appropriate for the target audience. Potential clinical participants (cardiologists, cardiac physiologists, and nurses) were identified through snowball sampling conducted by the principal investigators and clinical staff from each site, then invited to participate either via email or in person. Patients were identified through convenience sampling by site clinicians and were invited to participate via phone call or in person. There were no dropouts, and all participants who were approached agreed to partake in the study. Participants consented either electronically or verbally prior to study commencement. All interviews were conducted either over telephone calls or in person at a CIED clinic with only the researcher present. The interview duration ranged from 15 to 45 minutes. We continued to conduct interviews until the researcher judged that the dataset was sufficiently rich to meaningfully address the research question, conducting 35 interviews in total. This number exceeds the sample adequacy range suggested by Hennick and Kaiser, supporting the sufficiency of our sample. Interviews were audio recorded and transcribed verbatim, without field notes being taken. Participants did not receive a copy of the transcript to review or provide feedback on study findings.

### Data Analysis

Interview transcripts were uploaded to NVivo (version 14; QRS International Pty Ltd) software. Two investigators (BS and EO) analyzed using a hybrid approach, combining the benefits of an inductive thematic analysis with a deductive approach [14] to represent the data in a generalizable way using the TDF (Figure 1).

**Figure 1.** Thematic analysis of the interview data using an inductive and deductive approach. TDF: theoretical domains framework.




Data were analyzed in an iterative process. Initially 2 researchers (BS and EO) read and reread the first 5 transcripts and coding fragments relevant to the research question. The codes were reviewed, discussed, and deliberated between investigators (BS and EO) to compare the data interpretation. The deliberation aimed to ensure we had comprehensively covered all aspects of the research question, to explore any potential nuances in the interpretation, and resulted in the initial codebook development. One investigator (BS) continued the analysis of the remaining transcripts. This process was continually reviewed with refined versions of the codebook reviewed by the investigator (EO). This process enabled a transparent and rigorous approach to coding while remaining sensitive to the inductive and interpretive nature of the analysis.

Using a deductive analysis approach, codes were then matched to the appropriate TDF domains. This process was reviewed, discussed, and deliberated between investigators (BS and EO) until consensus was reached and consistent.

One investigator (BS) used an inductive analysis approach to develop subthemes from the codes before developing overarching themes [15]. Themes and subthemes were generated from codes across all participants, rather than stratifying by stakeholder title (cardiologists, cardiac physiologists, nurses, and patients). This process was reviewed and discussed between investigators (BS and EO) until a consensus was reached, resulting in the final data output.

### Trustworthiness

In the inductive phase, we ensured rigor by using structured codebooks and multiple coders to independently code the same data. The coders discussed their interpretations to refine and align them, ensuring consistency in the analysis while preserving the interpretive flexibility of the approach. In contrast, in the deductive phase, we applied the TDF to categorize themes. To ensure consistency and coherence in this process, we compared interpretations and reached a consensus on domain alignment. This collaborative approach helped enhance the reliability of our deductive analysis while respecting the interpretive nature of qualitative research.

### Ethical Considerations

Ethics approval was granted by the Western Sydney Local Health District (2022/ETH00271). All participants provided informed consent to partake in the study prior to data collection and were informed that they could withdraw from the study at any time. Participants were assigned a study ID and had all data deidentified. No form of compensation was provided to any participant for their involvement in the study.

## Results

### Overview

A total of 35 interviews were conducted between July 2022 and April 2023. In total, 16 of the interviews were conducted with patients, 10 with cardiologists, and 9 with cardiac physiologists and cardiac nurses. The mean patient age was 73.1 (SD 10.7)

years, and the majority were male (n=12, 75%) and born in Australia (n=12, 75%). Pacemakers (n=8, 50%) were the most common CIED type, and the mean duration of RM was 4.3 (SD 2.6) years. The mean cardiologist age was 46.2 (SD 6.3) years, and the majority were male (n=9, 90%), subspecialized in electrophysiology (n=7, 70%), had a mean duration of 12.3 (SD 6.6) years as a cardiologist, and a mean duration of 7.7 (3.6) years managing patients with RM. The mean physiologist or nurse age was 36.6 (SD 9.4), and the majority were female (n=5, 56%), and had a mean duration of 4.3 (SD 2.6) years managing patients with RM. Participant demographic and clinical experience results are presented in Table 1.

**Table .** Demographic, CIED[a], and RM[b] characteristics of interviewed stakeholders.

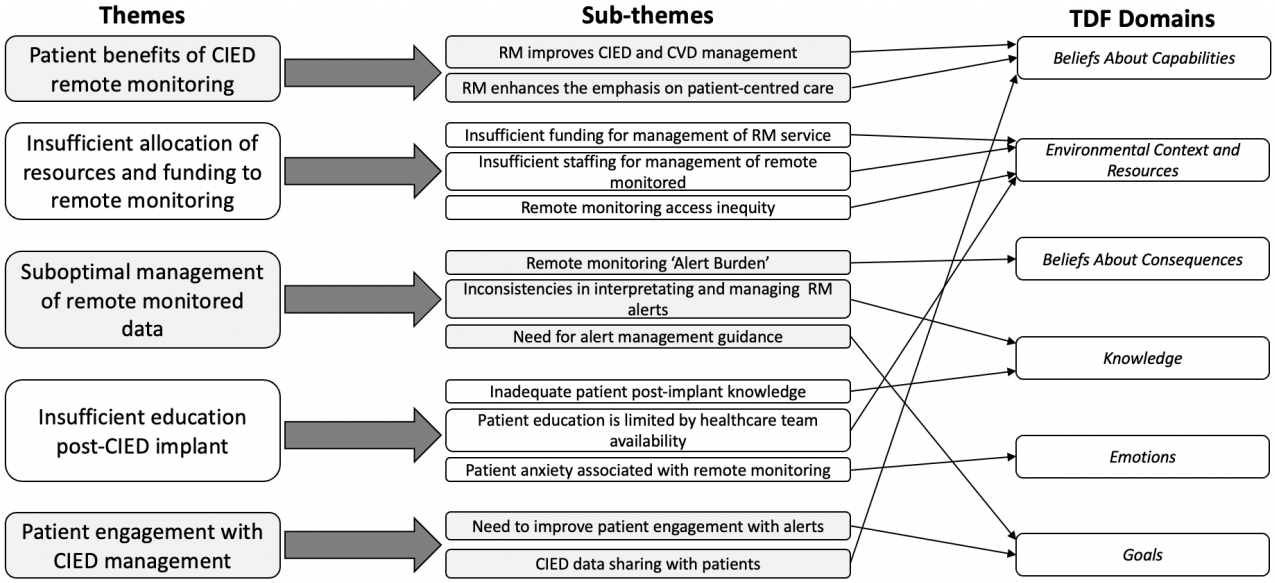
| Characteristic | Value (n=35) |
|---|---|
| Patients (n=16) | |
| Age (years), mean (SD) | 73.1 (10.7) |
| Male, n (%) | 12 (75) |
| Country of birth, n (%) | |
| Australia | 12 (75) |
| England | 3 (19) |
| Lebanon | 1 (6) |
| CIED indication, n (%) | |
| Ventricular tachycardia primary prevention | 7 (44) |
| Atrial fibrillation | 3 (19) |
| Bradycardia | 3 (19) |
| Syncope | 1 (6) |
| Arrhythmia (unknown to the patient) | 2 (12) |
| CIED type, n (%) | |
| Pacemaker | 8 (50) |
| Defibrillator | 5 (31) |
| Cardiac resynchronization therapy—pacemaker | 3 (19) |
| Duration receiving RM (years), mean (SD) | 4.3 (2.6) |
| Physiologists or nurses (n=9) | |
| Age (years), mean (SD) | 36.6 (9.4) |
| Male, n (%) | 4 (44) |
| Location, n (%) | |
| Western Sydney | 5 (56) |
| Illawarra | 1 (11) |
| Newcastle | 2 (22) |
| Sydney | 1 (11) |
| Duration managing RM (years), mean (SD) | 6.1 (2.6) |
| Cardiologists (n=10) | |
| Age (years), mean (SD) | 46.2 (6.3) |
| Male, n (%) | 9 (90) |
| Location, n (%) | |
| Western Sydney | 3 (30) |
| Illawarra | 1 (1) |
| Newcastle | 3 (3) |
| Sydney | 3 (3) |
| Cardiologist subspecialty, n (%) | |
| Electrophysiologist | 7 (70) |
| Heart failure specialist | 2 (20) |
| Proceduralist | 1 (10) |
| Duration as cardiologist (years), mean (SD) | 12.3 (6.6) |
| Duration managing RM (years), mean (SD) | 7.7 (3.6) |

[a]CIED: cardiac implantable electronic device.

[b]RM: remote monitoring.

We organized our results into themes and subthemes. Themes and subthemes are summarized in Figure 2, with subthemes and codes described below. One theme was deemed an enabler, and 4 barriers to RM. Illustrative quotes for each subtheme and code are presented in Tables S1-S13 in Multimedia Appendix 2.

**Figure 2.** Themes and subthemes emerged from qualitative thematic analysis with allocation to the relevant TDF domains. CIED: cardiac implantable electronic device; CVD: cardiovascular disease; TDF: theoretical domains framework.




## Theme 1: Patient Benefits on RM

### *RM Improves CIED and CVD Management*

The main benefits noted by stakeholders included the improved patient treatment outcomes facilitated by RM (Quotes 1-3 in Table S1 in Multimedia Appendix 2). These benefits were perceived to be largely driven by earlier detection of clinical issues (Quotes 4-9 in Table S1 in Multimedia Appendix 2), reduced postimplant issues (Quotes 10 and 11 in Table S1 in Multimedia Appendix 2), prevented hospital admissions (Quote 12 in Table S1 in Multimedia Appendix 2), and deployment of a service to rural and remote patients who otherwise have restricted access to CIED care (Quote 13 in Table S1 in Multimedia Appendix 2). Furthermore, clinicians reported that RM-based care enabled CIED management to be provided to patients without face-to-face review during the COVID-19 pandemic (Quote 14 in Table S1 in Multimedia Appendix 2).

### *RM Enhances the Emphasis on Patient-Centered Care*

Cardiologists noted RM processes are designed to be user-friendly for patients (Quote in Table S2 in Multimedia Appendix 2). Physiologists highlighted that RM facilitates improved care for patients in nursing homes, who previously had difficulties attending face-to-face clinics (Quote 2 in Table S2 in Multimedia Appendix 2). Patients expressed gratitude for the reduced hospital visits required for CIED reviews (Quote 3 in Table S2 in Multimedia Appendix 2). Additionally, patients reported a sense of safety derived from having the health care team monitor their data through RM (Quotes 4 and 5 in Table S2 in Multimedia Appendix 2).

## Theme 2: Insufficient RM Resources, Funding, and Recognition of Workload and Skills

### *Funding for Management of RM Service*

Barriers to the implementation and management of RM were centered around inadequate funding for clinics within the public sector. Cardiologists and physiologists reported that current reimbursement schemes fail to recognize the extensive tasks involved in providing the RM service and, in turn, do not provide adequate funding to deliver the service for improved patient care (Quotes 1-7 in Table S3 in Multimedia Appendix 2). Currently, clinicians reported that the delivery of RM comes with additional costs to the CIED clinics (Quote 8 in Table S3 in Multimedia Appendix 2), with some public hospitals reluctant to cover these costs despite the patient benefits (Quote 9 in Table S3 in Multimedia Appendix 2). Due to the inadequate funding, some clinicians reported they are unable to employ adequate staff to manage RM alerts (Quote 10 in Table S3 in Multimedia Appendix 2). Improved funding, infrastructure, and recognition by health services were recommended for RM development (Quotes 11 and 12 in Table S3 in Multimedia Appendix 2).

### *Staffing for Management of RM Alerts*

Interpreting and responding to alerts can be time-consuming due to the range of "invisible" tasks required, which include, but are not limited to, confirming the alert accuracy, reviewing previous alerts, patient history and medications, patient contact, reprogramming, education, report development, and cardiologist escalation. The time to complete these tasks varies among physiologists based on their experience and confidence levels (Quote 1 in Table S4 in Multimedia Appendix 2). Physiologists

mentioned that there is an inadequate number of staff employed to manage the RM workload (Quotes 2-4 in Table S10 in Multimedia Appendix 2), which can result in alerts not being managed in a timely fashion (Quotes 5 in Table S4 in Multimedia Appendix 2). Cardiologists mentioned that physiologists need more allocated time to manage alerts and scheduled reviews (Quotes 6 in Table S4 in Multimedia Appendix 2). Additionally, some cardiologists reported not having the capacity to review RM alerts (Quote 7 in Table S4 in Multimedia Appendix 2).

### *RM Access Inequity*

Not all patients receive RM, and factors associated with receiving RM drive inequity in access (Quotes 1 and 2 in Table S5 in Multimedia Appendix 2). Cardiologists identified some of these factors: public payment models are poorly suited to the provision of RM, existing health services may not provision RM support, and smaller services may not have the skill mix to support RM (Quotes 3 and 4 in Table S5 in Multimedia Appendix 2). Additionally, smaller cardiology clinics often lack the necessary resources and capacity to offer the service (Quote 5 in Table S5 in Multimedia Appendix 2). Other factors hindering the equitable distribution of RM include the incompatibility of CIED, with many older models unable to support this technology (Quote 6 in Table S5 in Multimedia Appendix 2), and inadequate patient internet access, particularly affecting rural patients (Quote 7 in Table S5 in Multimedia Appendix 2).

## Theme 3: RM Data Management Burden and Risks

### *RM "Alert Burden"*

Reviewing and managing alerts transmitted through RM was reported to be a time-consuming process for CIED clinic staff due to the range of "invisible" clinical and nonclinical tasks associated with alert receipt (Quotes 1-3 in Table S6 in Multimedia Appendix 2). Physiologists partly attributed this alert burden to their inability to modify alert parameters due to manufacturer system restrictions (Quotes 4 and 5 in Table S6 in Multimedia Appendix 2). Additionally, the transmission of alerts that are false positives further amplifies the workload for physiologists, which will only worsen with increasing CIED implants and RM utilization (Quote 6 in Table S6 in Multimedia Appendix 2). Consequently, the heightened alert burden resulting from a generalized alert setup and increased workload may compromise patient care and raise the likelihood of overlooking critical alerts (Quotes 7 and 8 in Table S6 in Multimedia Appendix 2).

### *Inconsistencies in Interpreting and Managing RM Alerts*

Physiologists raised that there is a lack of uniformity in knowledge, skills, experience, and training to manage RM alerts (Quotes 1-3 in Table S7 in Multimedia Appendix 2). It was noted that in some countries, the cardiac physiologist workforce regulation requires registration with a Clinical Physiologists Registration Board, but in other countries like Australia, this is not mandatory. It was also raised that the lack of more specific clinical guidelines, or pragmatic training on responding and managing RM alerts, presents risks and challenges to service delivery (Quotes 4 and 5 in Table S7 in Multimedia Appendix 2). Further participants highlighted that there were significant differences among cardiologists and cardiology services in the appropriate management of RM alerts (Quote 6 in Table S7 in Multimedia Appendix 2), including what information is relevant to convey by physiologists to clinicians upon alert detection (Quote 7 in Table S7 in Multimedia Appendix 2). Furthermore, cardiologists highlighted a lack of standardization in the "baseline" settings of alert thresholds (Quote 8 in Table S7 in Multimedia Appendix 2).

### *Need for Alert Management Guidance*

To enhance RM data management efficiency, physiologists have emphasized the need for RM alert management guidelines to provide support to CIED clinic staff (Quotes 1 and 2 in Table S8 in Multimedia Appendix 2). Additionally, cardiologists emphasize the importance of eliminating nonessential activities and implementing a process to receive alerts only for relevant, actionable issues (Quote 3 in Table S8 in Multimedia Appendix 2). Furthermore, cardiologists have expressed the need for a national consensus statement from experts in the RM field to provide standardized care for alert management (Quote 4 in Table S8 in Multimedia Appendix 2). Some clinics have taken the initiative to develop their internal alert management protocols, resulting in a reduction of "unnecessary" alerts and an overall decrease in workload (Quotes 5 and 6 in Table S8 in Multimedia Appendix 2).

## Theme 4: Insufficient Patient Education and Understanding of CIED and RM

### *Inadequate Patient Postimplant Knowledge*

Patients mentioned that the information provided post-CIED implant was inadequate for their needs. Key areas of knowledge deficit upon discharge included a poor understanding of the RM service (Quote 1 in Table S9 in Multimedia Appendix 2) and a poor understanding of restrictions to daily activities (Quotes 2-7 in Table S9 in Multimedia Appendix 2). A barrier to effective patient education can be the timing of information delivery, with patients reporting being overwhelmed peri-implant and struggling to retain information (Quote 8 in Table S9 in Multimedia Appendix 2). In addition, discrepancies in information delivery exist between CIED types, with physiologists reporting that patients with implantable cardioverter defibrillator routinely receive greater education than patients with permanent pacemaker (Quote 9 in Table S9 in Multimedia Appendix 2). Furthermore, discrepancies exist based on insurance status, with private patients often receiving greater information than public patients (Quotes 10 and 11 in Table S9 in Multimedia Appendix 2). Following hospital discharge, patients reported that there is a lack of resources to acquire information (Quote 12 in Table S9 in Multimedia Appendix 2) and a lack of communication channels to ask specific questions (Quote 13 in Table S9 in Multimedia Appendix 2). Ultimately, both patients and physiologists acknowledge that there is no formal postdischarge program available to provide ongoing patient education and support, which in the future is something that is required for RM progression (Quotes 14-17 in Table S9 in Multimedia Appendix 2).

#### *Patient Education is Limited by Health Care Team Availability*

Cardiologists acknowledged that discussions with patients and the delivery of "proper" education do not often occur, largely due to workload and time constraints (Quotes 1-2 in Table S10 in Multimedia Appendix 2). Both patients and physiologists believe insufficient explanations and education are provided to patients upon scheduled reviews (Quotes 3 and 4 in Table S10 in Multimedia Appendix 2). Patients frequently mentioned that they often have questions regarding their care and restrictions; however, they do not have access to the health care team to ask these questions (Quotes 5 and 6 in Table S10 in Multimedia Appendix 2).

#### *Patient Anxiety Associated With RM*

The use of RM could be associated with heightened patient anxiety, influenced by various factors. Cardiologists noted that patients may be hesitant to embrace the RM service, primarily due to concerns about the privacy of their data (Quote 1 in Table S11 in Multimedia Appendix 2). Patients reported that they experienced increased anxiety when receiving inconsistent information regarding their data, such as the battery life of their CIED (Quote 2 in Table S11 in Multimedia Appendix 2). In addition, patients reported that travel-related scenarios would exacerbate their anxiety, with patients and their families expressing mistrust in both the CIED and the RM system when traveling and not having close access to a hospital (Quotes 3 and 4 in Table S11 in Multimedia Appendix 2). This mistrust has stemmed from inconsistencies in patient explanations of CIED clinic and RM capabilities.

### Theme 5: Patient Engagement

#### *Need to Improve Patient Engagement With Alerts*

Patients and cardiologists mentioned the need for improved communication with patients following alert detection (Quotes 1-3 in Table S12 in Multimedia Appendix 2). However, patient contact should only occur if the alerts are actionable and relevant to the patient (Quotes 4 and 5 in Table S12 in Multimedia Appendix 2). Patients and physiologists mentioned the benefit of using a digital tool such as an SMS text messaging platform or app to contact patients regarding alerts and for patients to ask questions (Quotes 6-8 in Table S12 in Multimedia Appendix 2).

#### *CIED Data Sharing With Patients*

There were varying perspectives on the provision of CIED data to patients. Cardiologists felt that patients should be able to access their CIED data (Quote 1 in Table S13 in Multimedia Appendix 2) and that personalized in-time data provided to the patient would improve engagement (Quotes 2 and 3 in Table S13 in Multimedia Appendix 2). However, nurses and physiologists anticipate that data sharing could increase patient anxiety and concern (Quotes 4-6 in Table S13 in Multimedia Appendix 2). Patients noted that if they were to have access to their data, there would have to be careful consideration of what was presented (Quote 7 in Table S13 in Multimedia Appendix 2), and suggested that the data would need to be delivered in a user-friendly format (Quotes 8 and 9 in Table S13 in Multimedia Appendix 2).

### TDF

Subthemes were categorized into 6 TDF domains. The subthemes "RM improves CIED and CVD management," "RM enhances the emphasis on patient-centered care," and "CIED data sharing with patients" were developed within the Beliefs About Capabilities domain. Subthemes "Insufficient funding for management of RM service" "Insufficient staffing for management of remote monitored," and "Remote monitoring access inequity" were developed within the Environmental Context and Resources domain. The subtheme "Remote monitoring alert burden" was developed within the Beliefs About Consequences domain. Subthemes "Inconsistencies in interpreting and managing RM alerts" and "Inadequate patient postimplant knowledge" were developed within the Knowledge domain. The subtheme "Patient anxiety associated with remote monitoring" was developed within the Emotions domain. Finally, the subthemes "Need for alert management guidance" and "Need to improve patient engagement with alerts" were developed within the Goals domain.

## Discussion

### Principal Results

RM of CIEDs offers significant advantages for individuals with CVD; however, there is still a large scope for improved implementation. This study provides a current multidisciplinary perspective on RM implementation and a framework of barriers and enablers to address for improving future implementation and scale-up. We identified 5 main themes representing the barriers and facilitators to CIED with RM use. These themes are mapped to 6 domains of the TDF, which can inform targeted interventions to enhance implementation and maximize the potential benefits of CIED RM.

### Comparison With Other Work

Across the themes, there was a reinforcement of the benefits of CIED RM directly to the patient in both improved efficiencies in health care delivery and improved health outcomes through early detection of issues, prevention of hospital admissions, and better provision of care to rural or remote patients. These perspectives are corroborated by several recent studies which have demonstrated that RM enables earlier detection of actionable alerts [16], improves outcomes including reduced inappropriate shocks [17], decreases rates of strokes [16], and reduces mortality rates demonstrated in the pooled analysis of 3 RCTs using continuous RM [18]. Furthermore, improvements in health care service utilization have been demonstrated with reduced emergency department presentations [19], hospital admissions [20,21], and hospitalization length-of-stay times [21]. However, in patients with heart failure, RM has not consistently demonstrated benefits in mortality and heart failure hospital readmissions [22].

In total, 3 of the 5 themes identified centered on patient engagement, understanding, and perceived utility. Across subthemes, it was identified that RM enhances the focus on patient-centered care (offering a user-friendly service,

minimizing in-person reviews, correlating concerns with CIED data, and extending the service to patients who would otherwise lack such care) and enhances the patient's sense of care. This is underscored by expressions of patient satisfaction, appreciation, reassurance, and an improved sense of safety in managing their CIED and CVD. These observations align with prior studies that have consistently shown positive outcomes in terms of patient satisfaction [23,24], acceptance [25], and an enhanced feeling of safety [23,24,26].

However, resourcing and an inadequate recognition of the tasks arising from RM, as well as the skills and training needed to manage alerts, were consistently identified as barriers to CIED RM. Lack of funding and appropriate reimbursement schemes have also been seen as a prominent barrier in European and North American countries [6,27]. While a recent meta-analysis has demonstrated that CIED RM is a cost-effective intervention for health care systems [28], current models of care do not yet account for the additional tasks that arise from RM implementation, particularly those associated with alert management. Staff described alert management as comprising multiple additional phone calls, troubleshooting connectivity issues, alert triage, and scheduling in-person reviews [29]. Many staff and health services are not recognized for the increased workload associated with RM [27], which may be expected to rise with the increasing complexity of CVD, the complexity of technology, and the number of CIED implants.

RM data management was also consistently identified as a challenge to RM implementation. The "alert burden" associated with nonclinically significant alerts was particularly called out as a process management challenge. Contributing to this was the generalized nature of alert parameters, the discrepancies between alert interpretation, and the lack of clinical appropriateness guidance. Potential risks could also arise if the "alert burden" arising from "nonactionable" alerts jeopardizes patient care through the missing of time-critical alerts, a phenomenon described as "alert fatigue" [30]. Consequently, clinicians have expressed the need for the standardization of RM data management from the guidance of a national expert consensus panel. This call for RM standardization processes is not novel to this study, with multiple recent studies identifying the growing alert burden and need for guidance on standardized improved management approaches [6,31,32]. Recently, an international expert consensus statement was created by the Heart Rhythm Society and other large cardiac organizations to provide guidance for device clinics and clinicians on managing CIED follow-up, with some recommendations on operationalizing RM follow-up; however, this guidance lacks specificity on how to react to clinical issues detected via CIED RM [1]. In this study, some clinicians reported that their respective hospitals had instituted internal protocols for managing RM data, yielding positive outcomes in workload management without compromising patient care. Given the clinician's desire and potential benefits of a standardized approach to RM data management, improved clinical guidance on RM data management is required.

Insufficient post-CIED implant education was a key barrier identified across stakeholders. This study identified that many patients believe they do not receive adequate information, both peri-implant and upon discharge. This is in line with previous studies that have identified that patients have a substantial deficit in their CIED knowledge, despite having a strong desire to receive more information, specifically around restrictions on daily living and how to deal with device-related issues [8,9,33]. Clinicians noted that limited understanding of the technology by the patients can prevent the uptake of the RM service and increase patient anxiety living with a CIED. Despite this concern, clinicians noted that patient education is not enforced nor standardized, with variation seen in the provision of information due to factors such as CIED type, insurance status, CIED manufacturer, and clinic staff availability. Large language models show potential in addressing gaps in patient education for general cardiac risk factors [34]; however, further training is needed before clinicians can trust their ability to enhance understanding and engagement for patients with CIED [35]. Future co-design studies with key stakeholders are required to develop an effective and efficient program to allow adequate and standardized patient education, without significantly increasing clinician workload.

Finally, patient engagement with CIED management emerged as a prominent theme across stakeholders. CIED RM has the unique opportunity to better engage patients with their CVD management through the frequent transmission of cardiac data. Clinicians outlined that a future goal for RM is to better engage patients with the alerts received, through early contact on "actionable alerts." A potential modality proposed by stakeholders for this engagement is through a digital tool such as an SMS text messaging platform or app, where patients could access their data or alerts and communicate with their health care team. Clinicians had mixed beliefs on the utility of data sharing with patients, with some believing that it would positively increase engagement, while others are concerned it would increase patient anxiety and clinic workload. Patients believe that if data or alerts were to be provided to them, it would need to be presented in a user-friendly format. Previous studies focusing on CIED RM data interoperability with patients found that the data shared should be simplified, yet informative [36], be personalized and accompanied with informational support [37], and can ultimately enhance shared decision-making without increasing clinical workload [38]. Whilst CIED data sharing with patients may improve patient management, the feasibility of this technology is yet to be thoroughly explored.

## Strengths and Limitations

The strength of this study is the involvement of both patients and multidisciplinary clinicians, thus providing a comprehensive perspective of CIED RM barriers and enablers. The study also mapped the elicited themes and subthemes to behavior change techniques, which can be used to target actionable strategies for future adaptations to improve the RM service. However, this study has some limitations that need to be considered. First, participants were only recruited from New South Wales, Australia, with most included patients located in metropolitan and regional areas. However, the included multidisciplinary clinicians also serve patients from rural and remote regions and thus have a strong understanding of the barriers and enablers of the RM service in these areas. Second, the approach to participant recruitment used convenience sampling, which may

limit the generalizability of our results. Despite this, the participant population sampled is varied in their backgrounds, with patients having a wide spread of CIED types and indications for CIED implants, and clinicians having an appropriate mix of genders, occupations, and subspecializations for cardiologists. Thereby, the collected information is insightful and likely applicable to the wider population when informing future research and clinical directions of RM.

## Conclusions

This study highlights the benefits and challenges of CIED RM from the perspectives of patients and multidisciplinary clinicians. It emphasizes both the role of the patient with themes centering on patient engagement, education, and benefits, as well as that of multidisciplinary clinicians challenged by the wealth of data, alert burden, and complexity of tasks arising from RM. The findings can serve as a roadmap to action to guide the continued development and implementation of RM services into the future. It seems clear that there is great potential for patient and health system benefits from the implementation of good systems for RM, but we are not there yet.

## Acknowledgments

The authors would like to acknowledge the participants who were involved in the semistructured interview process. This work was supported by Biotronik Pty Ltd and the Australian Stroke and Heart Research Accelerator partnership project grant (grant G214388). AS is supported by an Australian Heart Foundation Future Leader Fellowship (ID106025). Generative artificial intelligence was not used in any portion of the manuscript generation.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

## Authors' Contributions

BS and CC conceptualized the research topic. BS, CC, AL, SL, and AS developed the methodology. BS recruited the patients. BS, CC, ETO, and KC acquired funding, curated the data, wrote, edited, and reviewed the original draft. All authors reviewed the final draft. CC provided supervision and accepts responsibility as the paper guarantor.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant interview guides.
[PDF File, 945 KB - cardio_v9i1e67758_app1.pdf ]

Multimedia Appendix 2
Subthemes, codes, and participant quotes.
[DOCX File, 43 KB - cardio_v9i1e67758_app2.docx ]

## References

1. Ferrick AM, Raj SR, Deneke T, et al. 2023 HRS/EHRA/APHRS/LAHRS expert consensus statement on practical management of the remote device clinic. J Arrhythm 2023 Jun;39(3):250-302. [doi: 10.1002/joa3.12851] [Medline: 37324757]
2. Leitch J, Asakai H, Dawson L, et al. Cardiac Society of Australia and New Zealand (CSANZ) position statement on the follow-up of cardiovascular implantable electronic devices 2022. Heart Lung Circ 2022 Aug;31(8):1054-1063. [doi: 10.1016/j.hlc.2022.05.001] [Medline: 35760743]
3. Lappegård KT, Moe F. Remote monitoring of CIEDs-for both safety, economy and convenience? Int J Environ Res Public Health 2021 Dec 28;19(1):312. [doi: 10.3390/ijerph19010312] [Medline: 35010572]
4. Varma N, Braunschweig F, Burri H, et al. Remote monitoring of cardiac implantable electronic devices and disease management. Europace 2023 Aug 2;25(9). [doi: 10.1093/europace/euad233] [Medline: 36857597]
5. Seiler A, Biundo E, Di Bacco M, et al. Clinic Time required for remote and in-person management of patients with cardiac devices: time and motion workflow evaluation. JMIR Cardio 2021 Oct 15;5(2):e27720. [doi: 10.2196/27720] [Medline: 34156344]
6. Kelly SE, Campbell D, Duhn LJ, et al. Remote monitoring of cardiovascular implantable electronic devices in Canada: survey of patients and device health care professionals. CJC Open 2021 Apr;3(4):391-399. [doi: 10.1016/j.cjco.2020.11.010] [Medline: 34027341]
7. Vandenberk B, Raj SR. Remote patient monitoring: what have we learned and where are we going? Curr Cardiovasc Risk Rep 2023;17(6):103-115. [doi: 10.1007/s12170-023-00720-7] [Medline: 37305214]

8. Patel D, Hu P, Hilow H, et al. The gap between what patients know and desire to learn about their cardiac implantable electronic devices. Pacing Clin Electrophysiol 2020 Jan;43(1):118-122. [doi: 10.1111/pace.13850] [Medline: 31782195]
9. Ghahari RR, Holden RJ, Flanagan ME, et al. Using cardiac implantable electronic device data to facilitate health decision making: a design study. Int J Ind Ergon 2018 Mar;64:143-154. [doi: 10.1016/j.ergon.2017.11.002]
10. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003 Oct 11;362(9391):1225-1230. [doi: 10.1016/S0140-6736(03)14546-1] [Medline: 14568747]
11. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
12. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37. [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
13. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
14. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
15. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006 Mar;5(1):80-92. [doi: 10.1177/160940690600500107]
16. Jang JP, Lin HT, Chen YJ, Hsieh MH, Huang YC. Role of remote monitoring in detection of atrial arrhythmia, stroke reduction, and use of anticoagulation therapy—a systematic review and meta-analysis. Circ J 2020 Oct 23;84(11):1922-1930. [doi: 10.1253/circj.CJ-20-0633] [Medline: 33012748]
17. Guédon-Moreau L, Lacroix D, Sadoul N, et al. A randomized study of remote follow-up of implantable cardioverter defibrillators: safety and efficacy report of the ECOST trial. Eur Heart J 2013 Feb;34(8):605-614. [doi: 10.1093/eurheartj/ehs425] [Medline: 23242192]
18. Hindricks G, Varma N, Kacet S, et al. Daily remote monitoring of implantable cardioverter-defibrillators: insights from the pooled patient-level data from three randomized controlled trials (IN-TIME, ECOST, TRUST). Eur Heart J 2017 Jun 7;38(22):1749-1755. [doi: 10.1093/eurheartj/ehx015] [Medline: 29688304]
19. Landolina M, Perego GB, Lunati M, et al. Remote monitoring reduces healthcare use and improves quality of care in heart failure patients with implantable defibrillators: the evolution of management strategies of heart failure patients with implantable defibrillators (EVOLVO) study. Circulation 2012 Jun 19;125(24):2985-2992. [doi: 10.1161/CIRCULATIONAHA.111.088971] [Medline: 22626743]
20. Mabo P, Victor F, Bazin P, et al. A randomized trial of long-term remote monitoring of pacemaker recipients (the COMPAS trial). Eur Heart J 2012 May;33(9):1105-1111. [doi: 10.1093/eurheartj/ehr419] [Medline: 22127418]
21. Piccini JP, Mittal S, Snell J, Prillinger JB, Dalal N, Varma N. Impact of remote monitoring on clinical events and associated health care utilization: a nationwide assessment. Heart Rhythm 2016 Dec;13(12):2279-2286. [doi: 10.1016/j.hrthm.2016.08.024] [Medline: 27544748]
22. McGee MJ, Ray M, Brienesse SC, et al. Remote monitoring in patients with heart failure with cardiac implantable electronic devices: a systematic review and meta-analysis. Open Heart 2022 Nov;9(2):e002096. [doi: 10.1136/openhrt-2022-002096] [Medline: 36442906]
23. Timmermans I, Meine M, Szendey I, et al. Remote monitoring of implantable cardioverter defibrillators: patient experiences and preferences for follow-up. Pacing Clin Electrophysiol 2019 Feb;42(2):120-129. [doi: 10.1111/pace.13574] [Medline: 30536931]
24. Artico J, Zecchin M, Zorzin Fantasia A, et al. Long-term patient satisfaction with implanted device remote monitoring: a comparison among different systems. J Cardiovasc Med (Hagerstown) 2019 Aug;20(8):542-550. [doi: 10.2459/JCM.0000000000000818] [Medline: 31107287]
25. Ricci RP, Morichelli L, Quarta L, et al. Long-term patient acceptance of and satisfaction with implanted device remote monitoring. Europace 2010 May;12(5):674-679. [doi: 10.1093/europace/euq046] [Medline: 20200019]
26. Maciąg A, Mitkowski P, Mazurek M, et al. Patient perspective and safety of remote monitoring of implantable cardioverter-defibrillators in the Polish Nationwide Multicenter Registry: the Medtronic CareLink network evaluation. Kardiol Pol 2020 Nov 25;78(11):1115-1121. [doi: 10.33963/KP.15556] [Medline: 32794684]
27. Mairesse GH, Braunschweig F, Klersy K, Cowie MR, Leyva F. Implementation and reimbursement of remote monitoring for cardiac implantable electronic devices in Europe: a survey from the health economics committee of the European Heart Rhythm Association. EP Europace 2015 May;17(5):814-818. [doi: 10.1093/europace/euu390]
28. Arcinas LA, Alyosif M, Rennert-May E, et al. Cost-effectiveness of remote monitoring for cardiac implantable electronic devices compared with conventional follow-up: a systematic review. Eur Heart J 2023 Nov 9;44(Supplement_2). [doi: 10.1093/eurheartj/ehad655.696]
29. Ryan P, McGrath C, Lawrie I, Fitzsimons C, O'Shea J, De BrÚn J. Enhancing efficiency in a cardiac investigations department by increasing remote patient monitoring. Int J Qual Health Care 2019 Dec 22;31(Supplement_1):29-34. [doi: 10.1093/intqhc/mzz065] [Medline: 31867661]

30. Alert fatigue. Patient Safety Network. 2019. URL: https://psnet.ahrq.gov/primer/alert-fatigue#:~:text=The%20term%20alert%20fatigue%20describes,respond%20appropriately%20to%20such%20warnings [accessed 2024-10-09]
31. O'Shea CJ, Middeldorp ME, Hendriks JM, et al. Remote monitoring alert burden: an analysis of transmission in >26,000 patients. JACC Clin Electrophysiol 2021 Feb;7(2):226-234. [doi: 10.1016/j.jacep.2020.08.029] [Medline: 33602404]
32. Harvey M, Seiler A. Challenges in managing a remote monitoring device clinic. Heart Rhythm O2 2022 Feb;3(1):3-7. [doi: 10.1016/j.hroo.2021.12.002] [Medline: 35243430]
33. Góral S, Teliżyn M, Rajzer M, Olszanecka A. Patient's knowledge of daily activities, need for information and quality of life after cardiac electronic device implantation. Folia Med Cracov 2022 Jun 29;62(1):121-134. [doi: 10.24425/fmc.2022.141695] [Medline: 36088597]
34. O'Hagan E, McIntyre D, Laranjo L. The potential for a chat-based artificial intelligence model to facilitate educational messaging on hypertension. Hypertension 2023 Aug;80(8):e128-e130. [doi: 10.1161/HYPERTENSIONAHA.123.21395] [Medline: 37325936]
35. Hillmann HAK, Angelini E, Karfoul N, Feickert S, Mueller-Leisse J, Duncker D. Accuracy and comprehensibility of chat-based artificial intelligence for patient information on atrial fibrillation and cardiac implantable electronic devices. Europace 2023 Dec 28;26(1). [doi: 10.1093/europace/euad369]
36. Daley C, Ghahari RR, Drouin M, et al. Involving patients as key stakeholders in the design of cardiovascular implantable electronic device data dashboards: Implications for patient care. Heart Rhythm O2 2020 Jun;1(2):136-146. [doi: 10.1016/j.hroo.2020.04.005] [Medline: 34113868]
37. Toscos T, Daley C, Wagner S, et al. Patient responses to daily cardiac resynchronization therapy device data: a pilot trial assessing a novel patient-centered digital dashboard in everyday life. Cardiovasc Digital Health J 2020;1(2):97-106. [doi: 10.1016/j.cvdhj.2020.09.003] [Medline: 35265880]
38. Mirro M, Daley C, Wagner S, Rohani Ghahari R, Drouin M, Toscos T. Delivering remote monitoring data to patients with implantable cardioverter-defibrillators: does medium matter? Pacing Clin Electrophysiol 2018 Nov;41(11):1526-1535. [doi: 10.1111/pace.13505] [Medline: 30225880]

## Abbreviations

**CIED:** cardiac implantable electronic device
**COREQ:** Consolidated Criteria for Reporting Qualitative Research
**CVD:** cardiovascular disease
**RM:** remote monitoring
**TDF:** theoretical domains framework

*Edited by A Coristine; submitted 20.10.24; peer-reviewed by E Scruth, M Nomali; revised version received 01.05.25; accepted 01.05.25; published 18.07.25.*

*Please cite as:*
*Sheahen B, O'Hagan ET, Cho K, Shaw T, Lee A, Lal S, Sverdlov AL, Chow C*
*Barriers and Enablers to Routine Clinical Implementation of Cardiac Implantable Electronic Device Remote Monitoring in Australia Among Cardiologists, Cardiac Physiologists, Nurses, and Patients: Interview Study*
*JMIR Cardio 2025;9:e67758*
*URL: https://cardio.jmir.org/2025/1/e67758*
*doi:10.2196/67758*

# Telecardiology Activities in Hospital and University Cardiology Facilities in Italy: Survey Study

Manuela Bocchino[1,2*], MD; Elvira Agazio[1*], MD; Cecilia Damiano[1*], PhD; Giuseppe Di Lorenzo[3*], MD; Fabio De Paolis[1]; Duilio Luca Bacocco[4], MSc; Fabrizio Ammirati[5], MD; Alessio Nardini[6], MSc; Marco Silano[1*], MD

[1]Department of Cardiovascular, Endocrine-Metabolic Diseases and Aging, Istituto Superiore di Sanità, Roma, Italy

[2]Head of the Telemedicine UOS, General Management Staff, ASL Roma 3, Via Casal Bernocchi, 73, Roma, Italy

[3]School of Specialization in Hygiene and Public Health, Sapienza University of Rome, Rome, Italy

[4]Office for General Affairs, Istituto Superiore di Sanità, Roma, Italy

[5]Emergency Department, Division of Cardiology, Azienda Sanitaria Locale Roma 3, Rome, Italy

[6]Ministero della Salute, Rome, Italy

[*]these authors contributed equally

**Corresponding Author:**
Manuela Bocchino, MD
Department of Cardiovascular, Endocrine-Metabolic Diseases and Aging, Istituto Superiore di Sanità, Roma, Italy

## Abstract

**Background:** Telemedicine enables the provision of health services at a distance using information and communication technologies and includes different types of services: telemonitoring, remote control, virtual visit or televisit, telereferral, teleassistance, medical teleconsultation, health professionals' teleconsultation, and telerehabilitation. Continuous monitoring, early care, and greater therapeutic adherence could be benefits of telemedicine in the management of cardiovascular diseases. There are not many studies in the literature investigating the use of telemedicine in cardiology in Italy.

**Objective:** The aim of this study is to illustrate the results of a survey on telemedicine services in cardiology conducted by the Department of Cardiovascular, Endocrine-Metabolic Diseases and Aging of the Italian National Institute of Health.

**Methods:** The Telehealth Quality of Care Tool (TQoCT) from the World Health Organization (WHO) was used as the model. A survey was disseminated by the National Association of Doctors and Hospital Cardiologists (ANMCO) from June 2024 to October 2024 through a link provided to hospital and university cardiology operative units identified through the 8th Census of Cardiological Structures in Italy. The facilities were contacted by email or telephone. The survey was built using Microsoft Forms and composed of 52 questions divided into 6 sections. The analysis was carried out for the whole national territory and by geographical area.

**Results:** Of the 443 hospitals contacted, the response rate was 56.7% (251/443). Overall, 78.9% (198/251) of facilities reported telemedicine initiatives providing telemonitoring (128/198, 64.6%), telereferrals (104/198, 52.5%), medical teleconsultations (93/198, 47%), televisits (82/198, 41.4%), health professionals' teleconsultations (64/198, 32.3%), and telerehabilitation (10/198, 5.1%). The most frequently followed cardiovascular conditions were heart failure, ischemic heart disease, and cardiac arrhythmias, especially atrial fibrillation. Of the facilities, 51% (101/198) used deliberations, procedures, protocols, or informed consent for their activities, and 46% (91/198) of the reported services were paid. Lack of dedicated staff, complexity in organizational terms, and lack of technological equipment in the structure were the principal obstacles for health professionals; lack of familiarity with technology was the principal obstacle for patients.

**Conclusions:** There are still organizational and clinical limitations to resolve to make telemedicine in cardiology an integral part of medical practice. The true challenge of telecardiology is likely the integration of available technology with precise, concrete, and simplified organizational models. As a tool, technology is fundamental only if it is accessible and adequate. However, it must be integrated with new paths built according to the needs of the territory, patients, and health personnel. Such a survey could provide help for the future design and use of telemedicine services in cardiology in Italy.

***(JMIR Cardio 2025;9:e73747)*** doi:10.2196/73747

**KEYWORDS**

telemedicine; cardiology; survey; methodology; design

## Introduction

### Background

According to the definition from the World Health Organization (WHO), telemedicine is the "provision of health services by all actors using information and communication technologies (ICT) for the exchange of valid information for diagnosis, treatment and prevention of diseases where distance is a critical factor, in the interest of improving the health of individuals and their communities" [1]. Telemedicine therefore allows remote health services to be delivered using ICT devices without patients having to go to a health care facility. Telemedicine includes different types of services, such as telemonitoring, remote control, virtual visit or televisit, telereferral, teleassistance, medical teleconsultation, health professionals' teleconsultations, and telerehabilitation.

Regarding cardiology, telemedicine can be a useful tool for postdischarge care continuity, for example. Cardiology is a branch with a considerable number of chronic diseases, so continuous monitoring can help prevent and manage the worsening of a patient's condition. Similarly, it allows remote data sharing, providing a comparison among health care professionals in a context that is not possible onsite.

In Italy, according to the mapping of experiences in telemedicine for the national territory drawn up in 2018 [2], telemonitoring and teleassistance represented 26% of the services, with 43% concerning cardiology specialists and 57% representing other specialists. Many regions and autonomous provinces took note of the 2020 national guidelines or interim indications for telemedicine care services [3,4] in different documents that proposed and implemented local solutions. In the field of cardiology, the "National Consensus Document on telemedicine for cardiovascular diseases: indications for tele rehabilitation and telemonitoring" [5] was recently published. Although telecardiology proved beneficial in terms of care, comfort, and patient access, it still has limitations, both logistical and medical, to becoming an integral part of medical practice in the future [6].

In 2021, the Italian Association of Arrhythmology and Cardiostimulation (AIAC) released a survey entitled "Telecardiology and the use of remote monitoring of implantable Devices in Italy in light of the COVID-19 period." That survey aimed mainly to investigate the spread of remote monitoring of cardiac implantable electronic devices (CIEDs) in Italy, in particular following the period of the health emergency [7]. Subsequently, in October 2023, the Italian National Institute of Health (ISS) collaborated on the "First national survey on telemedicine in private outpatient care" with Luiss University, the Health Wellness and Resilience Observatory of the Bruno Visentini Foundation, and the integrative health fund Fasdac [8]. This survey concerned the use of telemedicine within private outpatient facilities. Recently, the 8th Census of Cardiological Structures in Italy, conducted by the National Association of Doctors and Hospital Cardiologists (ANMCO), reported some data about the spread of telemedicine in cardiology in Italy, but it only covered the frequency of use of televisit services, telemonitoring, teleconsultations, and telereferrals. The census was conducted in 2022 and published in the beginning of 2024 [9]. At the international level, the literature covers surveys on perceptions of health professionals about telemedicine during the COVID-19 pandemic [10] and the acceptance of telemedicine in cardiology by specialists and general practitioners [11]. Currently, there are no dedicated surveys reported in the literature that investigated telemedicine services and their characteristics in cardiology in Italy.

### Aim of the Study

This study reports the results of the survey on telecardiology in Italy, which was conceived, designed, and conducted by the Department of Cardiovascular, Endocrine-Metabolic Diseases and Aging of the ISS. The survey was addressed to hospital and university cardiology operative units (OUs) in the national territory—both those listed by ANMCO and those found using an online search.

The aim was to describe the current state of telecardiology services in Italy: Being aware of real-world data on the territory is essential to be able to understand the strengths and weaknesses of a service as well as to improve it.

## Methods

This report was conducted in adherence with CHERRIES (Checklist for Reporting Results of Internet E-Surveys) [12]. The Telehealth Quality of Care Tool (TQoCT) [13] was also consulted. Based on previous recommendations and scoreboards, the WHO developed a tool that serves both as a guide and as a self-assessment to improve the quality of telemedicine services, which is useful at the local, regional, and national levels.

### Recruitment

The respondents were selected on the basis of the 8th Census of Cardiological Structures in Italy conducted by ANMCO [9] in 2022. We also performed an online research of the institutional sites of the Italian regional health service for a better understanding of the real operational structures.

The questionnaire was disseminated in 3 levels of intervention: email with a cover letter sent to the health departments of the regions and autonomous provinces, email with a cover letter addressed to the Directors of the hospital and university cardiology OUs, and phone contact with the Directors of the cardiology hospital and university OUs.

An ISS email account was specifically created for the survey (surveytelecardiologia@iss.it) as well as the cover letters.

The institutional emails of the Directors of the OUs were used as contact details, unless they had explicitly requested to send them to their personal email address; the publicly available phone numbers of the OUs or numbers voluntarily provided by Directors themselves were used for telephone contact. The survey started on June 14, 2024, and ended on October 18, 2024.

To ensure widespread territorial coverage, several scientific societies were involved in survey dissemination, by sending emails with a cover letter to societies' Presidents. In particular, the Italian Society of Cardiology; the AIAC, especially the Women & Arrhythmology Area; the Latium section of the

Italian Society of Telemedicine; and the Italian Society of Digital Medicine actively contributed to the study.

A Microsoft Forms template was designed to ensure that the IP addresses of the computers used to complete the survey were not recorded. For each participating facility, the Microsoft Forms application created a unique ID and collected information on start time and completion time of data entry. The data controller was ISS as well as the participating centers.

### The Questionnaire

The survey, digitally composed using Microsoft Forms, was based on a questionnaire of 52 questions divided into 6 sections: (1) facility information (6 questions); (2) presence of telemedicine projects (1 question); (3) type of telemedicine services (27 questions) defined according to the 2020 State-Regions Agreement [4]; (4) governance, safety, and accessibility (12 questions); (5) volumes and patient centrality (4 questions); and (6) obstacles to service development and implementation (2 questions).

Of the 52 total questions, 18 were multiple choice, and 11 included a text description field, 4 of which were dedicated to the description of the structure. In addition, 16 required a single answer, while 6 required date entries and 1 question required numerical data. The complete questionnaire is available in Multimedia Appendix 1.

### Statistical Analysis

Statistical elaboration of the collected data was performed using R software (version 4.0.3) both on a national level and by geographical area. In particular, regions and autonomous provinces were divided into the 3 macro-areas of “north” (Aosta Valley, Piedmont, Liguria, Lombardy, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, and Emilia-Romagna), “center” (Tuscany, Marche, Umbria, Latium, and Abruzzo), and “south and islands” (Molise, Campania, Puglia, Basilicata, Calabria, Sicily, and Sardinia). Quantitative variables are expressed as n (%).

### Ethical Considerations

The survey was submitted to the National Ethics Committee for Trials of Public Research Institutions and other public bodies with a national character (National Ethics Committee), which approved the study on May 17, 2024 (protocol number 0021567). The activity did not collect personal data; in any case, the study information and consent to participate in the study were an integral part of the form accessed through the link shared to participate in the survey itself.

## Results

### Facility Information

Of 443 facilities contacted, 251 responded, for a response rate of 56.7%. Table 1 shows the distribution of contacted and responding centers, along with their response rates, reported by region or autonomous province, and geographical areas.

**Table .** Contacted centers and responding centers reported by region or autonomous province and geographical area.

| Location | Contacted centers (n=443), n (%) | Responding centers (n=251), n (%)[a] |
|---|---|---|
| Region | | |
| Abruzzo | 13 (2.9) | 0 (0) |
| Basilicata | 4 (0.9) | 1 (25) |
| Calabria | 14 (3.2) | 7 (50) |
| Campania | 46 (10.4) | 12 (26.1) |
| Emilia-Romagna | 29 (6.5) | 19 (65.5) |
| Friuli-Venezia Giulia | 8 (1.8) | 5 (62.5) |
| Latium | 47 (10.6) | 36 (76.6) |
| Liguria | 19 (4.3) | 13 (68.4) |
| Lombardy | 60 (13.5) | 36 (60.0) |
| Marche | 13 (2.9) | 5 (38.5) |
| Molise | 4 (0.9) | 1 (25) |
| Piedmont | 31 (7) | 25 (80.6) |
| Puglia | 29 (6.5) | 14 (48.3) |
| Sardinia | 10 (2.3) | 9 (90) |
| Sicily | 41 (9.3) | 19 (46.3) |
| Tuscany | 30 (6.8) | 22 (73.3) |
| Trentino-Alto Adige | 5 (1.1) | 3 (60) |
| Umbria | 8 (1.8) | 5 (62.5) |
| Aosta Valley | 1 (0.2) | 1 (100) |
| Veneto | 31 (7) | 18 (58.1) |
| Geographical area | | |
| North | 184 (41.5) | 120 (65.2) |
| Center | 98 (22.1) | 68 (69.4) |
| South and islands | 161 (36.3) | 63 (39.1) |

[a]Percentage of those who were contacted in that region.

The mean time to complete the survey was 12 (SD 5) minutes. About one-half (123/251, 49%) of the responding facilities were part of hospitals directly managed by local health companies, followed by hospital and university companies (Table S1 in Multimedia Appendix 2).

## Presence of Telemedicine Projects

In total, 198 facilities (198/251, 78.9%) reported telemedicine initiatives; details by geographical areas are shown in Table 2.

**Table .** Facilities using telemedicine by geographical area.

| Use of telemedicine | North (n=120), n (%) | Center (n=68), n (%) | South and islands (n=63), n (%) | Total (n=251), n (%) |
|---|---|---|---|---|
| Yes | 101 (84.2) | 53 (77.9) | 44 (69.8) | 198 (78.9) |
| No | 19 (15.8) | 15 (22.1) | 19 (30.2) | 53 (21.1) |

## Type of Telemedicine Services

Since the same facility could perform more than just one telecardiology service, respondents were able to answer this question with multiple options. We found that the majority of facilities (128/198, 64.6%) provided telemonitoring services, followed by telereferrals (104/198, 52.5%), medical teleconsultations (93/198, 47%), televisits (82/198, 41.4%), health professionals' teleconsultations (64/198, 32.3%), and telerehabilitation (10/198, 5.1%). Table 3 shows the breakdown by geographical area.

**Table .** Telemedicine services provided by facilities using telemedicine, reported by geographical area.

| Telemedicine service | North (n=101), n (%) | Center (n=53), n (%) | South and islands (n=44), n (%) | Total (n=198, n (%) |
|---|---|---|---|---|
| Televisit | 50 (49.5) | 24 (45.3) | 8 (18.2) | 82 (41.4) |
| Medical teleconsultation | 52 (51.5) | 28 (52.8) | 13 (29.6) | 93 (47) |
| Health professionals' teleconsultation | 31 (30.7) | 21 (39.6) | 12 (27.3) | 64 (32.3) |
| Telemonitoring | 73 (72.3) | 34 (64.2) | 21 (47.7) | 128 (64.6) |
| Telereferral | 54 (53.5) | 27 (50.9) | 23 (52.3) | 104 (52.5) |
| Telerehabilitation | 8 (7.9) | 2 (3.8) | 0 (0) | 10 (5.1) |

Table 4 reports the cardiovascular diseases that were most commonly followed through different telemedicine services: patients with heart failure, carriers of CIEDs, or patients who needed therapeutic plans related to cardiovascular diseases.

**Table .** Cardiovascular diseases for which a national telemedicine service was provided in facilities using telemedicine (n=198).

| Disease | Televisit (n=82), n (%) | Medical teleconsultation (n=93), n (%) | Health professionals' teleconsultation (n=64), n (%) | Telemonitoring (n=128), n (%) | Telerehabilitation (n=10), n (%) |
|---|---|---|---|---|---|
| Ischemic heart disease | 19 (23.2) | 51 (54.8) | 16 (25) | 6 (4.7) | 5 (50) |
| Congenital heart disease | 1 (1.2) | 16 (17.2) | 1 (1.6) | —[a] | 1 (10) |
| Valvular heart disease | 3 (3.7) | 32 (34.4) | 6 (9.4) | 1 (0.8) | 3 (30) |
| Remote control of CIED[b] | 55 (67.1) | 36 (38.7) | 37 (57.8) | 90 (70.3) | — |
| Atrial fibrillation | 20 (24.4) | 39 (41.9) | 16 (25) | 46 (35.9) | — |
| Arterial hypertension | 8 (9.8) | 21 (22.6) | 6 (9.4) | 5 (3.9) | — |
| Rare diseases in cardiology | 8 (9.8) | 14 (15.1) | 3 (4.7) | 3 (2.3) | 1 (10) |
| Arrhythmic disorders | 15 (18.3) | 35 (37.6) | 17 (26.6) | 41 (32) | 2 (20) |
| Treatment plans for cardiovascular diseases | 40 (48.8) | — | — | — | — |
| Noncardiological diseases | — | 12 (12.9) | 4 (6.2) | — | — |
| Heart failure | 64 (78) | 64 (68.8) | 40 (62.5) | 69 (53.9) | 8 (80) |
| Syncope | 8 (9.8) | 26 (28) | 14 (21.9) | 41 (32) | — |

[a]Not applicable.

[b]CIED: cardiac implantable electronic device.

A televisit was often used to follow up heart failure, for the remote control of CIEDs, and for the compilation of therapeutic plans of the drugs monitored by the Italian Agency of the Drug. Medical teleconsultations and health professionals' teleconsultations were associated mainly with heart failure. Telemonitoring was mainly used with patients with CIEDs and to a lesser extent with patients using wearable devices, or at least not implantable devices, who had heart failure and atrial fibrillation. Finally, heart failure and ischemic heart disease were the pathologies for which telerehabilitation was used most often.

Telereferral of cardiological instrumental examinations was primarily used for the remote interpretation of electrocardiograms (ECGs), particularly within hospital-to-territory, intrahospital, and emergency network settings. Further details are shown in Table 5.

**Table .** Services that involved telereferrals at the national level and by geographical area.

| Service | North (n=54), n (%) | Center (n=27), n (%) | South and islands (n=23), n (%) | Total (n=104), n (%) |
|---|---|---|---|---|
| Remote control of CIED[a] | 34 (63) | 14 (51.8) | 8 (34.8) | 56 (53.8) |
| Emergency ECG[b] (network 118) | 38 (70.4) | 9 (33.3) | 15 (65.2) | 62 (59.6) |
| Intrahospital ECG | 41 (75.9) | 19 (70.4) | 21 (91.3) | 81 (77.9) |
| Hospital-territory ECG | 42 (77.8) | 13 (48.1) | 7 (30.4) | 62 (59.6) |
| Echocardiogram | 5 (9.3) | 1 (3.7) | 3 (13) | 9 (8.7) |
| ECG Holter monitoring for 24 hours | 9 (16.7) | 8 (29.6) | 5 (21.7) | 22 (21.8) |
| ECG monitoring through other devices | 7 (13) | 3 (11.1) | 3 (13) | 13 (12.5) |
| Outpatient monitoring of blood pressure | 1 (1.8) | 6 (22.2) | 3 (13) | 10 (9.6) |

[a]CIED: cardiac implantable electronic device.

[b]ECG: electrocardiogram.

For some telemedicine services, additional elements were investigated (ie, the facilities [hubs and spoke] in which the medical teleconsultation was carried out and number and skills of the health personnel involved in teleconsultation and telemonitoring).

Medical teleconsultations were carried out mainly with other regional hospitals and with general practitioners or free choice pediatricians, followed by the OUs of the same hospital (Table 6). There was still little interaction between regions, although this was present at greater proportions in the regions of central Italy; in the south, at a territorial level, interactions with territorial cardiology prevailed rather than with general medicine.

**Table .** Facilities in which medical teleconsultation was carried out at the national level and by geographical area.

| Type of structure | North (n=52), n (%) | Center (n=28), n (%) | South and islands (n=13), n (%) | Total (n=93), n (%) |
|---|---|---|---|---|
| National hospitals or referral centers for pathology | 7 (14) | 5 (18) | 1 (8) | 13 (14) |
| Regional hospitals | 27 (52) | 13 (46) | 6 (46) | 46 (50) |
| Territorial cardiology clinics | 7 (14) | 2 (7) | 6 (46) | 15 (16) |
| Specialist noncardiological territorial | 4 (8) | 1 (4) | 0 (0) | 5 (5) |
| General practitioners or free choice pediatricians | 27 (52) | 5 (18) | 3 (23) | 35 (38) |
| Private or private/contracted facilities | 5 (10) | 1 (4) | 3 (23) | 9 (10) |
| OUs[a] of the same hospital | 9 (17) | 15 (54) | 9 (69) | 33 (36) |

[a]OU: operative unit.

Hub centers accounted for 30% (28/93), spoke centers accounted for 40% (37/93), and the remaining worked as both hub and spoke centers (details by geographical area are shown in Table S4 in Multimedia Appendix 2).

The health professionals' teleconsultation service was primarily conducted by nurses, followed by cardiocirculatory pathophysiology technicians and, at lower percentages, doctors, physiotherapists, and bioengineers. Details can be found in Table 7.

**Table .** Staff providing health professionals' teleconsultation services at the national level and by geographical area.

| Professional role | North (n=31), n (%) | Center (n=21), n (%) | South and islands (n=12), n (%) | Total (n=64), n (%) |
|---|---|---|---|---|
| Bioengineers | 0 (0) | 2 (10) | 2 (17) | 4 (6) |
| Cardiologists | 0 (0) | 0 (0) | 2 (17) | 2 (3) |
| Physiotherapists | 1 (3) | 0 (0) | 1 (8) | 2 (3) |
| Nurses | 26 (84) | 13 (62) | 7 (58) | 46 (72) |
| Doctors | 4 (13) | 4 (19) | 3 (25) | 11 (17) |
| Doctors-radiology technicians | 1 (3) | 0 (0) | 0 (0) | 1 (2) |
| Cardiocirculatory pathophysiology technicians | 8 (26) | 6 (29) | 2 (17) | 16 (25) |

Telemonitoring was primarily managed by doctors and, to a lesser extent, by nurses and cardiocirculatory pathophysiology technicians (Table S3 in Multimedia Appendix 2). In particular, nursing staff performed more telemonitoring services in the north than in the center and south; cardiocirculatory pathophysiology technicians were more involved in the north and center, less in the south (19%, 4/21).

## Governance, Safety, and Accessibility

Only 46% (91/198) of the centers were subjected to reimbursement. The reimbursement mainly concerned televisits and remote control (ie, remote control of the CIED) and, to a lesser extent, telemonitoring, teleconsultations, and other services (Table 8).

**Table .** Presence of reimbursement for telemedicine services and service types at the national level and by geographical area.

| Reimbursement characteristics | North (n=101), n (%) | Center (n=53), n (%) | South and islands (n=44), n (%) | Total (n=198), n (%) |
|---|---|---|---|---|
| Charges | | | | |
| Yes | 56 (55.4) | 27 (50.9) | 8 (18.2) | 91 (46) |
| No | 45 (44.6) | 26 (49.1) | 36 (81.8) | 107 (54) |
| Services subject to charges | | | | |
| Renewal of treatment plans | 0 (0) | 0 (0) | 1 (12.5) | 1 (1.1) |
| Health professionals' teleconsultations | 5 (10.9) | 2 (7.4) | 0 (0) | 7 (7.7) |
| Medical teleconsultations | 8 (14.3) | 4 (57.1) | 0 (0) | 12 (13.2) |
| Remote control (CIED[a]) | 32 (57.1) | 10 (37) | 5 (62.5) | 47 (51.6) |
| Telemonitoring | 16 (28.6) | 2 (7.4) | 2 (25) | 20 (22) |
| Telerehabilitation | 4 (7.1) | 0 (0) | 0 (0) | 4 (4.4) |
| Televisits | 31 (55.3) | 23 (85.2) | 5 (62.5) | 59 (64.8) |
| I don't know | 1 (1.8) | 1 (3.7) | 0 (0) | 2 (2.2) |

[a]CIED: cardiac implantable electronic device.

Considering the facilities that provided telemedicine services, in 51% of cases (101/198), company deliberations, specific procedures, operational protocols, or informed consent are used. In 43% of the cases (86/198), the structure did not prepare any of the aforementioned documents for the personnel addresses (Table 9). For more detailed information on the type of documents adopted, please consult Table S4 in Multimedia Appendix 2.

**Table .** In the facilities using telemedicine, presence of deliberations, procedures, protocols, or informed consent at the national level and by geographical area.

| Presence of the documents | North (n=101), n (%) | Center (n=53), n (%) | South and islands (n=44), n (%) | Total (n=198), n (%) |
|---|---|---|---|---|
| Yes | 57 (57.4) | 26 (49.1) | 17 (38.6) | 101 (51) |
| No | 40 (39.6) | 22 (41.5) | 24 (54.6) | 86 (43.4) |
| I don’t know | 2 (2) | 5 (9.4) | 0 (0) | 7 (3.6) |
| Other | 1 (1) | 0 (0) | 3 (6.8) | 4 (2) |

From a technological point of view, 52% (103/198) of the centers used devices (47/101, 46.5% in the north; 27/53, 51% in the center; and 28/44, 64% in the south), including medical devices in 86% (89/198; 88/101, 87.1% in the north; 47/53, 89% in the center; and 36/44, 82% in the south) of centers. Table 10 shows the device types used; for further details on device types, refer to Table S5 in Multimedia Appendix 2.

**Table .** Devices used in telemedicine facilities nationally and by geographical area.

| Device | North (n=101), n (%) | Center (n=53), n (%) | South and islands (n=44), n (%) | Total (n=198), n (%) |
|---|---|---|---|---|
| Implantable device[a] | 28 (27.7) | 15 (28.3) | 13 (29.5) | 56 (28.3) |
| Device for monitoring basic clinical parameters[b] | 59 (58.4) | 35 (66) | 44 (100) | 138 (69.7) |
| ECG[c] | 18 (17.8) | 12 (22.6) | 18 (41) | 48 (24.2) |
| Ultrasound | 0 (0) | 1 (1.9) | 1 (2.3) | 2 (1) |
| Event recorder | 0 (0) | 1 (1.9) | 0 (0) | 1 (0.5) |
| Spirometer | 3 (2.3) | 1 (1.9) | 0 (0) | 4 (2.0) |
| Other | 1 (1.0) | 0 (0) | 0 (0) | 1 (0.5) |

[a]Hemodynamic monitoring devices, loop recorder, pacemaker also biventricular, and defibrillators also biventricular.

[b]Physical activity, sleep, weight, heart rate, blood pressure, glycemia, body composition and hydration status of the body, pulse oximetry, temperature, and respirator frequency.

[c]ECG: electrocardiogram.

An interesting aspect concerned the health care workers’ access to telemedicine services. Only 56 of the 198 respondents (28.3%) had a telemedicine center with dedicated staff (33/101, 32.7% in the north; 13/53, 24% in the center; and 10/44, 23% in the south) in the facility. The facility or OU provided training for health personnel to start telemedicine services in 56% of cases (111/198; 64/101, 63.4% in the north; 24/53, 45% in the center; and 24/44, 54% in the south). Only 39 of the 198 facilities (19.7%) were able to collect data.

## Volume and Patient Centrality

The number of patients treated by most facilities ranged from 100 to 500 annually; details are given in Table 11.

**Table .** Patients followed yearly through telemedicine at the national level and by geographical area.

| Number of patients treated using telemedicine | North (n=101), n (%) | Center (n=53), n (%) | South and islands (n=44), n (%) | Total (n=198), n (%) |
|---|---|---|---|---|
| 0 - 100 | 13 (12.9) | 13 (24.6) | 11 (25) | 37 (18.7) |
| 101 - 500 | 62 (61.4) | 33 (62.2) | 29 (65.9) | 124 (62.6) |
| >500 | 26 (25.7) | 7 (13.2) | 4 (9.1) | 37 (18.7) |

The majority of facilities reported adequate information services provided to patients, and communication was mostly personalized at the start of the patients’ treatment (83/198, 41.9%); information services included informative brochures, face-to-face meetings, or digital material. In 25.3% of cases (50/198), none of these instruments were provided. The detail, at the national level and by geographical area, is shown in Table 12.

**Table .** Information tools for patients and their caregivers on telemedicine services and the use of necessary technology at the national level and by geographical area.

| Information tools | North (n=101), n (%) | Center (n=53), n (%) | South and islands (n=44), n (%) | Total (n=198), n (%) |
|---|---|---|---|---|
| In customized form at the beginning of the treatment pathway | 52 (51.7) | 23 (44.1) | 9 (20.6) | 83 (41.9) |
| Informative brochures | 24 (23.3) | 9 (16.2) | 6 (12.7) | 37 (18.7) |
| Events in attendance | 11 (10.8) | 2 (1.9) | 3 (6.3) | 15 (7.6) |
| Digital material on the web-site | 3 (3.3) | 2 (1.9) | 0 (0) | 5 (2.5) |
| None of these | 20 (19.8) | 13 (23.5) | 17 (38.1) | 50 (25.3) |

In only 20.2% (40/198) of the cases, a patient satisfaction assessment was carried out through questionnaires (27/40, 68%) or interviews (13/40, 32%).

## Obstacles to Service Development and Implementation

The obstacles to the development of telemedicine are reported in Table 13. The most reported problems were a lack of dedicated personnel, organizational complexity, and a lack of technological endowments available in the structure. Regulatory problems relating to the lack of clear and simple rules for all and technological barriers, such as poor connectivity and poor digital information of staff, were also present. Data about obstacles were similar between facilities not providing telemedicine services and those providing them.

**Table .** Obstacles encountered in the adoption of telemedicine services at the national level and by geographical area.

| Obstacles encountered | Geographical area | | | Use of telemedicine | |
|---|---|---|---|---|---|
| | North (n=101), n (%) | Center (n=53), n (%) | South and islands (n=44), n (%) | No (n=53), n (%) | Yes (n=198), n (%) |
| No obstacles | 5 (4.2) | 4 (5.9) | 2 (3.2) | 1 (1.8) | 10 (5) |
| Lack of legislation | 26 (21.7) | 15 (22.1) | 17 (27.0) | 7 (13.2) | 51 (25.7) |
| Complexity in organizational terms | 63 (52.5) | 37 (54.4) | 35 (55.6) | 23 (43.4) | 112 (56.6) |
| Complexity in the application of GDPR[a] | 15 (12.5) | 5 (7.4) | 9 (14.3) | 2 (3.8) | 27 (13.6) |
| Complexity in the application of specific legislation | 16 (13.3) | 4 (5.9) | 8 (12.7) | 7 (13.2) | 21 (10.6) |
| Mistrust and lack of confidence in telemedicine | 7 (5.8) | 3 (4.4) | 6 (9.5) | 2 (3.8) | 14 (7.1) |
| Data fragmentation and noninteroperability of systems | 26 (21.7) | 17 (25) | 17 (27) | 6 (11.3) | 54 (27.3) |
| Lack of technological equipment in the structure | 44 (36.7) | 34 (50) | 24 (38.1) | 23 (43.4) | 79 (39.9) |
| Lack of adequate infrastructure and Internet connectivity | 28 (23.3) | 17 (25) | 18 (28.6) | 8 (15.1) | 55 (27.8) |
| Lack of dedicated staff | 64 (53.3) | 38 (55.9) | 47 (74.6) | 30 (56.6) | 119 (60.1) |
| Cost in economic terms | 21 (17.5) | 7 (10.3) | 11 (17.5) | 8 (15.1) | 31 (15.6) |
| Internal procedures inappropriate to the theme | 19 (15.8) | 6 (8.8) | 11 (17.5) | 5 (9.4) | 31 (15.6) |
| Poor digital training of the staff | 12 (10) | 12 (17.6) | 10 (15.9) | 5 (9.4) | 29 (14.6) |
| Lack of staff willingness or cooperation | 6 (5) | 7 (10.3) | 6 (9.5) | 4 (7.5) | 15 (7.6) |

[a]GDPR: General Data Protection Regulation.

The obstacles to the development of telemedicine encountered by patients included a lack of familiarity with technology, followed by difficulty with using software and devices, lack of information, and poor home internet connection. Distrust toward these new services concerned only 35 of the 198 centers (17.7%). Details are shown in Figures 1 and 2.

**Figure 1.** Obstacles to the adoption of telemedicine tools by patients at the national level.

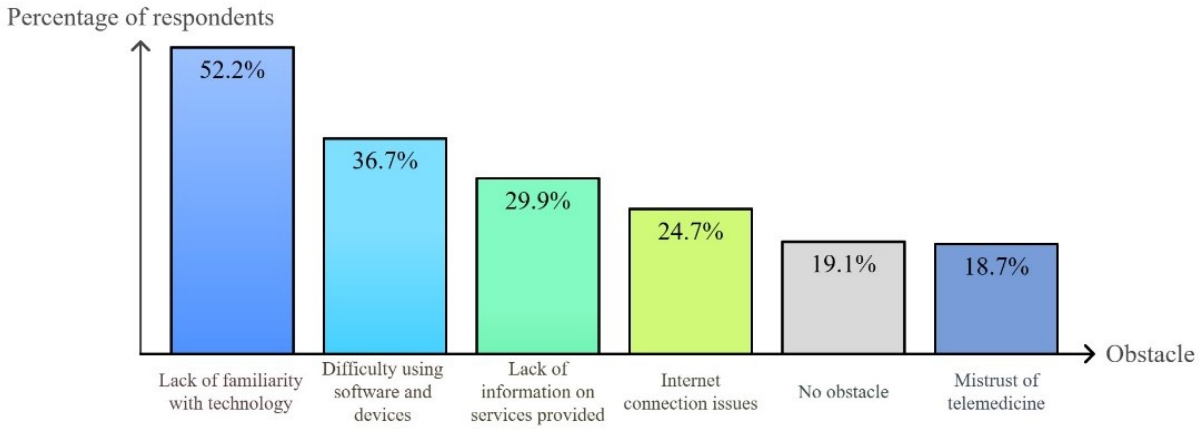
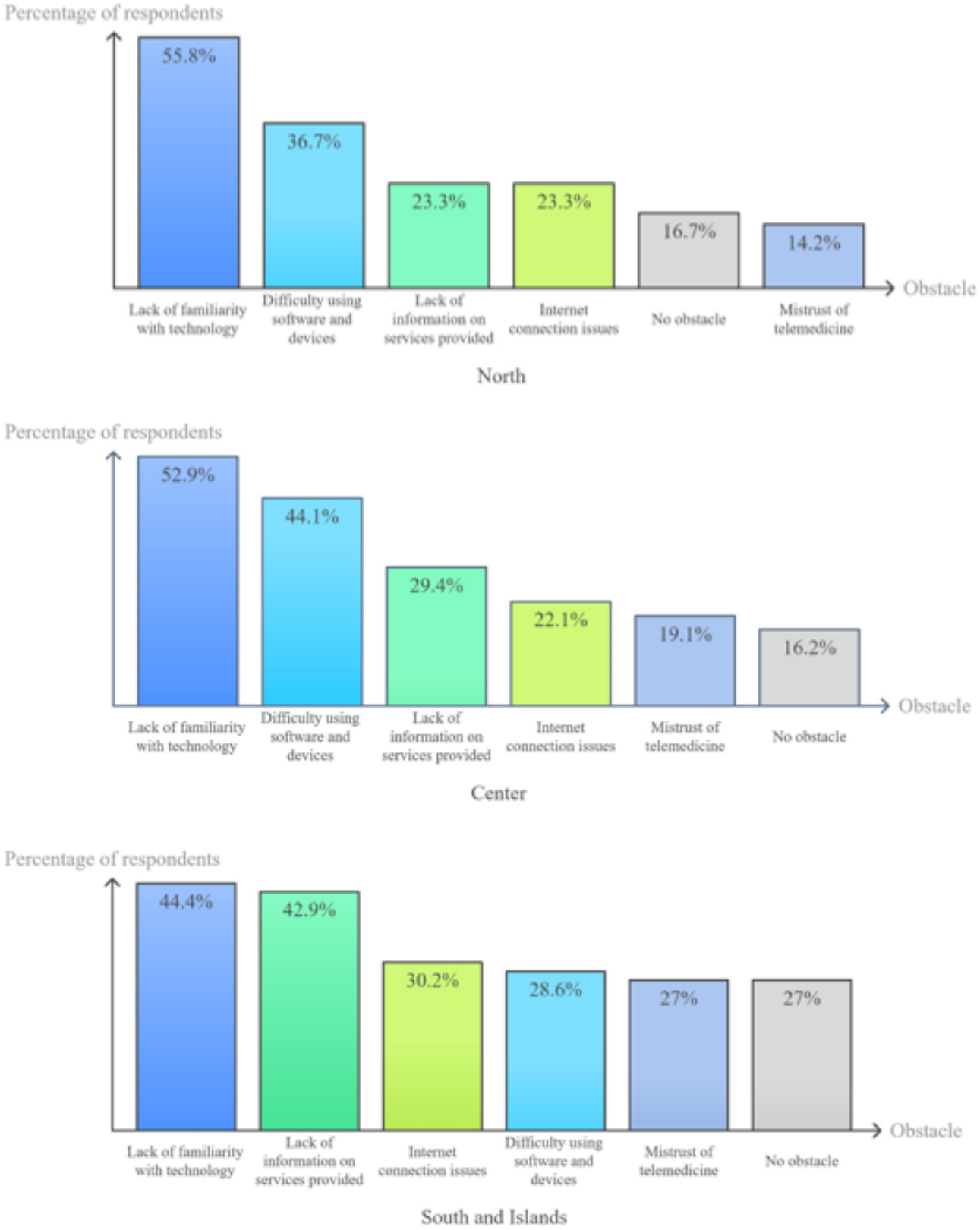


**Figure 2.** Barriers to patient uptake of telemedicine tools by geographical breakdown: north, center, and south and islands.




## Discussion

### Principal Findings

The response rate was in line with other published surveys, although not very high, and almost 80% of the centers reported telemedicine services in cardiology, most frequently telemonitoring, telereferrals, and medical teleconsultations. Heart failure was the disease for which these services were most frequently used. The possibility of reimbursement was present in less than one-half the cases; being guided by deliberations or operational protocols emerged as the situation for one-half the cases. The devices used in most facilities were medical devices, but health staff was rarely dedicated to these activities,

trained in just over 55% of these activities. In addition, only 20% of staff had the ability to collect data. In only one-quarter of cases, patients and caregivers were not adequately informed.

The major reported concerns included a lack of dedicated staff, organizational complexity, and a lack of technological endowments available in the structure. For patients, the following were concerns: a lack of familiarity with technology, difficulties using software and devices, a lack of information, and a poor home internet connection.

These survey results warrant several considerations. The first concerns the response rate (56.7%), which, although not optimal, allowed the information of interest to be obtained from a considerable number of centers. Connection, cooperation, and dialogue between institutional organizations, which monitor and carry out research, and operational structures remain fundamental aspects to understand the critical points and solve them to build paths adapted to real-world needs and achieve results.

The most widespread telemedicine services and diseases treated through these tools would suggest that telemedicine could provide better home care, especially for chronic cases. This would allow early control of signs of instability or poor adherence to therapy, facilitating interception at the initial stages of clinical deterioration and a possible timely change in therapeutic strategy when the patient is still asymptomatic, before the onset of the acute event. However, the data on the effectiveness and impact of such interventions available in the literature are still controversial.

A systematic review conducted in 2022 analyzed 72 studies with more than 127,000 participants, showing that remote monitoring and consultations for patients with heart failure were associated with a reduced risk of cardiovascular mortality and hospitalization [14]. However, a review conducted in 2021 concluded that the pooled effect estimate of telemedicine interventions on all‐cause hospitalizations and all‐cause deaths in patients with recently decompensated heart failure and without implanted devices was neutral [15]. Similarly, another systematic review in the same year demonstrated that most telemonitoring programs do not show clear effects on health care utilization measures, except for an increase in nonemergency outpatient department visits [16].

Furthermore, there is a clear demand for greater professional interaction, which could accelerate the diagnostic and therapeutic process. Articles on teleconsultations published in the last 10 years show that there are positive impacts of teleconsultation such as improving patient management, but there are still gaps that need to be addressed [17].

There are still significant unresolved issues that limit the spread of telemedicine services, not only in the cardiology field but also regarding the reimbursement of telemedicine services, for example. At the international level, a qualitative study in the United States established that staff perceive current telehealth reimbursement policies as a factor that exacerbates inequities to accessing care. These findings indicate that, although telehealth brings new opportunities to advance patient-centered care, there are serious challenges on the path toward equitable care because telehealth is not yet integrated into the payment system in a sustainable way, and this is a factor exacerbating the lack of staff [18]. In Europe, a paper published in April 2025 [19] presented a comparative analysis of remote patient monitoring in 7 European countries (France, Germany, Italy, the Netherlands, Poland, Spain, and the United Kingdom): Germany introduced specific reimbursement codes for remote monitoring of patients with heart failure as early as 2020; the Netherlands has a relatively more advanced and flexible refund system, with dedicated codes for different digital performances; and in the United Kingdom, reimbursement is relatively more developed, especially in regions with more advanced health systems. In France, reimbursement for remote monitoring is still very limited and partial; in Poland, it is very limited; and in Spain, it is still poorly structured and often supported by pilot projects or temporary funds. Finally, as shown, in Italy, reimbursement is fragmented at the regional level, with some regions further ahead and others still in the embryonic stage. The inadequacy of traditional reimbursement models and the need for sustainable, evidence-based methods are highlighted as necessary for real digital transformation, as well as to obtain dedicated staff, which represents a major obstacle to development [20].

Additional obstacles are the resulting organizational complexity, also linked to the lack of appropriate organizational models supported by the presence of procedures and operational protocols that address and protect health personnel and patients [19]. Last but not least, the lack of adequate technological equipment remains a critical issue.

The centrality of the patient is a fundamental aspect. The data collected suggest that properly informing patients about these new services and methodologies is a need. This is fundamental to achieving good therapeutic adherence, increasing patient empowerment and engagement, and helping the patient and caregiver understand how much it can improve quality of life and therefore increase therapeutic continuity and avoid possible related risks [21-23]. Despite some geographical variability, the relatively low percentage of mistrust toward these new services (19% nationally, 14% in the north, 16% in the center, and 27% in the south) suggests that much was invested in recent times for patient centrality and providing information to patients and caregivers, probably obtaining good outcomes. The results show that 25% of patients are still not properly informed; it is necessary to supplement these data with evidence that some of the services provided (telereporting of ECGs, for example, or certain instrumental examinations) do not require detailed information to the patient on the functioning of the service itself. Unfortunately, it was not possible to perform a subanalysis to check in detail whether this percentage of “inadequate” information was related to these services, since the question relating to patient information was related to the telecardiology service in general and, therefore, to all services. Apart from this possible correlation, this lack of information provided to the patient justifies the main obstacles identified: the low familiarity with and difficulty using the essential technologies and the lack of information on the services provided.

### Limitations

The survey had some limitations regarding planning and dissemination. Since it was not possible to disseminate the survey through regional health departments, we contacted the facilities directly as described in the Methods section; thus, there is the possibility that we failed to contact some structures. In addition, considering the 11 questions that included a text description field, 7 of them asked the respondents to specify the cardiovascular diseases treated using telemedicine services. Due to the variety and dispersion of responses, it was difficult to elaborate answers to those questions; therefore, we decided to limit the analysis only to major cardiovascular pathologies, without providing further details. A final consideration concerns the remote referral: This service was investigated only in relation to instrumental examinations (eg, ECG, echocardiogram, monitoring of heart rhythm and blood pressure) without considering other services, televisits, and teleconsultation for which telereferring is provided.

Additionally, we cannot exclude that facilities using telemedicine services were more likely to participate in the survey than those without a telemedicine service. Finally, as this was a survey-based study, answers to some questions may have been affected by inaccuracies due to self-report bias or interpretation of the questions. Fortunately, the percentage of this event was very low.

### Conclusions

There are still organizational and clinical limitations in the application of telemedicine in cardiology, one of the first fields to use these services in real-life patient management; telemedicine in cardiology is not yet an integral part of medical practice. In fact, it appears that the spread of its use in Italy is still limited because of some structural and organizational barriers. Mainly, there is a lack of dedicated staff to manage and implement telemedicine services; sometimes, there are no clear rules and regulations on the uniform and safe adoption of these technologies at the national level or information about them. Finally, despite their wide availability, technologies are usually not well integrated into care pathways. These results confirm that the true challenge of telecardiology is the integration of available technology with precise, concrete, and simplified organizational models. As a tool, technology is fundamental only if it is accessible and adequate, and it must be integrated with new paths built according to the needs of the territory, patients, and health personnel. Fundamental aspects are reporting procedures and operating protocols to establish who does what and how and the availability of dedicated, adequately trained personnel who are able to inform patients and caregivers.

Such research could, in the future, help the implementation of a national plan for telemedicine services in cardiology and help create a telemedicine network to support continuous improvement of national and regional strategies.

## Acknowledgments

We would like to thank the scientific societies that contributed to the survey dissemination: Italian Society of Cardiology (SIC), Italian Association of Arrhythmology and Cardiostimulation (AIAC), Women & Arrhythmology Area of the AIAC, Latium regional section of the Italian Society of Telemedicine (SIT), and Italian Society of Digital Medicine (SIMeDi). We also extend special thanks to the directors of the Cardiology Operating Unit who responded and gave their important contribution to the survey. The full list of participating health facilities is reported in Multimedia Appendix 3. It should be noted that the number of facilities does not correspond to the total number of responses received, since sometimes, there are several operating units within the same hospital.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Microsoft Forms questionnaire.
[DOCX File, 36 KB - cardio_v9i1e73747_app1.docx ]

Multimedia Appendix 2
Additional information about the hospitals, centers, personnel, and devices.
[DOCX File, 32 KB - cardio_v9i1e73747_app2.docx ]

Multimedia Appendix 3
List of health facilities who participated in the survey.
[DOCX File, 33 KB - cardio_v9i1e73747_app3.docx ]

## References

1. Telemedicine: opportunities and developments in member states. report on the second global survey on eHealth. World Health Organization. 2010. URL: https://iris.who.int/handle/10665/44497 [accessed 2025-11-16]

2. Mapping of telemedicine experiences on the territory in 2018 [Website in Italian]. Ministry of Health. 2018. URL: https://www.salute.gov.it/imgs/C_17_pagineAree_2515_2_file.pdf [accessed 2025-11-16]
3. Gabbrielli F, Bertinato L, De Filippis G, Bonomini M, Cipolla M. Interim indications for telemedicine care services during the COVID-19 health emergency. Report ISS COVID-19, no. 12/2020 [Website in Italian]. Italian National Institute of Health. 2020 Dec 15. URL: https://www.iss.it/rapporti?p_p_id=com_liferay_portal_search_web_portlet_SearchPortlet&p_p_lifecycle=0&p_p_state=maximized&p_p_mode=view&_com_liferay_portal_search_web_portlet_SearchPortlet_mvcPath=%2Fview_content.jsp&_com_liferay_portal_search_web_portlet_SearchPortlet_assetEntryId=5332135&_com_liferay_portal_search_web_portlet_SearchPortlet_type=content&p_l_back_url=https%3A%2F%2F [accessed 2025-11-18]
4. Permanent conference for the relations between the state, the regions and the autonomous provinces of Trento and Bolzano. Agreement, pursuant to article 4, paragraph 1, of the legislative decree 28 August 1997, n. 281, between the government, the regions and the autonomous provinces of Trento and Bolzano, the provinces, the municipalities and the mountain communities, on the document "National indications for the provision of services in telemedicine" [Website in Italian]. Presidency of the Council of Ministers. 2020 Dec 17. URL: https://www.statoregioni.it/media/3221/p-3-csr-rep-n-215-17dic2020.pdf [accessed 2025-11-16]
5. Gabbrielli F, Gruppo di Consensus Nazionale sulla Telecardiologia. Documento di consensus nazionale sullatelemedicina per le patologie cardiovascolari: indicazioni per la teleriabilitazione e il telemonitoraggio (Rapporti ISTISAN 23/21) [Report in Italian]. : Istituto Superioredi Sanità; 2023 URL: https://www.iss.it/documents/20126/9340234/23-21+web.pdf/edb447f0-19ba-71df-59c8-3db1ff88e696?t=1711719777176 [accessed 2025-11-25]
6. Mahalwar G, Kumar A, Kalra A. Virtual cardiology: past, present, future directions, and considerations. Curr Cardiovasc Risk Rep 2023;17(6):117-122. [doi: 10.1007/s12170-023-00719-0] [Medline: 37305213]
7. Maines M, Palmisano P, Del Greco M, et al. Impact of COVID-19 pandemic on remote monitoring of cardiac implantable electronic devices in Italy: results of a survey promoted by AIAC (Italian Association of Arrhythmology and Cardiac Pacing). J Clin Med 2021 Sep 10;10(18):4086. [doi: 10.3390/jcm10184086] [Medline: 34575197]
8. Health Wellbeing and Resilience Observatory. Connecting the dots: towards a national plan of health. Fondazione per la Ricerca Economica e Sociale ETS. 2023. URL: https://www.fondazioneries.it/wp-content/uploads/2025/01/Connecting-the-dots_-toward-a-national-plan-of-Health.-II-Report-2023.pdf [accessed 2024-12-19]
9. Oliva F, Di Pasquale G, Lucci D, et al. 8th census of cardiology centers in Italy. Italian Association of Hospital Cardiologists (ANMCO). Year 2022. G Ital Cardiol (Rome) 2024 Feb;25(2 Suppl 1):19S-103S. [doi: 10.1714/4195.41821] [Medline: 38291920]
10. Elawady A, Khalil A, Assaf O, Toure S, Cassidy C. Telemedicine during COVID-19: a survey of health care professionals' perceptions. Monaldi Arch Chest Dis 2020 Sep 22;90(4). [doi: 10.4081/monaldi.2020.1528] [Medline: 32959627]
11. Muehlensiepen F, Hoffmann MJ, Nübel J, et al. Acceptance of telemedicine by specialists and general practitioners in cardiology care: cross-sectional survey study. JMIR Form Res 2024 Feb 20;8(8):e49526. [doi: 10.2196/49526] [Medline: 38376898]
12. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
13. Telehealth quality of care tool. World Health Organization. 2024 Mar 13. URL: https://www.who.int/czechia/publications/i/item/WHO-EURO-2024-9475-49247-73556 [accessed 2024-12-19]
14. Kuan PX, Chan WK, Fern Ying DK, et al. Efficacy of telemedicine for the management of cardiovascular disease: a systematic review and meta-analysis. Lancet Digit Health 2022 Sep;4(9):e676-e691. [doi: 10.1016/S2589-7500(22)00124-8] [Medline: 36028290]
15. Drews TEI, Laukkanen J, Nieminen T. Non-invasive home telemonitoring in patients with decompensated heart failure: a systematic review and meta-analysis. ESC Heart Fail 2021 Oct;8(5):3696-3708. [doi: 10.1002/ehf2.13475] [Medline: 34165912]
16. Auener SL, Remers TEP, van Dulmen SA, Westert GP, Kool RB, Jeurissen PPT. The effect of noninvasive telemonitoring for chronic heart failure on health care utilization: systematic review. J Med Internet Res 2021 Sep 29;23(9):e26744. [doi: 10.2196/26744] [Medline: 34586072]
17. Deldar K, Bahaadinbeigy K, Tara SM. Teleconsultation and clinical decision making: a systematic review. Acta Inform Med 2016 Jul 16;24(4):286-292. [doi: 10.5455/aim.2016.24.286-292] [Medline: 27708494]
18. Porteny T, Brophy SA, Burroughs E. Experiences of telehealth reimbursement policies in federally qualified health centers. JAMA Netw Open 2025 Feb 3;8(2):e2459554. [doi: 10.1001/jamanetworkopen.2024.59554] [Medline: 39937474]
19. Helms TM, Boriani G, Brunner-La Rocca HP, et al. The present and future of cardiological telemonitoring in Europe: a statement from seven European countries. Herzschr Elektrophys 2025 Apr;36(2):138-148. [doi: 10.1007/s00399-025-01076-8]
20. Connolly J, Martin SS, Joynt Maddox KE, Maddox TM. Reimbursement transformation Is essential for digital health adoption in cardiovascular care. JAMA Cardiol 2025 Jul 1;10(7):641-642. [doi: 10.1001/jamacardio.2025.1013] [Medline: 40366677]
21. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 3;7(8):e016242. [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]

22. Nielsen MK, Johannessen H. Patient empowerment and involvement in telemedicine. J Nurs Educ Pract 2019;9(8):54. [doi: 10.5430/jnep.v9n8p54]
23. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020 Aug 12;8(8):e20359. [doi: 10.2196/20359] [Medline: 32784177]

## Abbreviations

**AIAC:** Italian Association of Arrhythmology and Cardiostimulation
**ANMCO:** National Association of Doctors and Hospital Cardiologists
**CHERRIES:** Checklist for Reporting Results of Internet E-Surveys
**CIED:** cardiac implantable electronic device
**ECG:** electrocardiogram
**ICT:** information and communication technology
**ISS:** Italian National Institute of Health
**OU:** operative unit
**TQoCT:** Telehealth Quality of Care Tool
**WHO:** World Health Organization

*Edited by J Rivers; submitted 12.03.25; peer-reviewed by M Maines, M Tanhapour; revised version received 08.07.25; accepted 15.07.25; published 05.12.25.*

*Please cite as:*
*Bocchino M, Agazio E, Damiano C, Di Lorenzo G, De Paolis F, Bacocco DL, Ammirati F, Nardini A, Silano M*
*Telecardiology Activities in Hospital and University Cardiology Facilities in Italy: Survey Study*
*JMIR Cardio 2025;9:e73747*
*URL: https://cardio.jmir.org/2025/1/e73747*
*doi:10.2196/73747*

# Patient Preferences for Using Remote Care Technology in Heart Failure: Discrete Choice Experiment

Ahmed Al-Naher[1], BSc, MBBS, PGCert, PhD; Jennifer Downing[2], PhD; Dyfrig Hughes[1,3], PhD; Munir Pirmohamed[1], ChB (Hons), MB, PhD

[1]Institute of Systems, Molecular and Integrative Biology (ISMIB), University of Liverpool, Liverpool, United Kingdom

[2]Institute of Population Health, University of Liverpool, Block B: Waterhouse Buildings, 1-5 Brownlow Street, Liverpool, United Kingdom

[3]Centre for Health Economics and Medicines Evaluation, Bangor University, Bangor, United Kingdom

**Corresponding Author:**
Jennifer Downing, PhD
Institute of Population Health, University of Liverpool, Block B: Waterhouse Buildings, 1-5 Brownlow Street, Liverpool, United Kingdom

## Abstract

**Background:** Remote care technology has been used to bridge the gap between health care in a clinical setting and in the community, all the more essential post-COVID. Patients with chronic conditions may benefit from interventions that could provide more continuous and frequent monitoring of their disease process and support self-management. A common barrier, however, is the lack of engagement with technological interventions or devices that provide care remotely, which could lead to loss of resources invested and reduced quality of care.

**Objective:** This discrete choice experiment elicits the preferences of patients with heart failure with regard to potential remote care technologies that they would be willing to engage with and, in turn, creates a hierarchy of factors that can affect engagement for use within future technology design.

**Methods:** A survey was created using discrete choice design and with input from a patient and public involvement group. It was distributed online via social media to patients with heart failure and to patient support groups. The attributes used for the experiment were based on a previous systematic review looking at factors that affect engagement in remote care and which generated five central themes, each of which was assigned to an attribute directly: communication (increasing interaction between patients and health care staff/carers/other patients), clinical care (improving the quality of care compared to established practice), education (providing tailored information to help with self-care and reduce uncertainty), ease of use (the technical aspects of the intervention are easy to handle without issues), and convenience (the intervention fits well around the patient's lifestyle and requires minimal effort). Each of the five themes had two levels, positive and negative. The survey presented participants with multiple forced-choice two-alternative scenarios of remote care, which allowed them to trade attributes according to their preference. The results were analyzed using binary logit to obtain preference weights for each attribute.

**Results:** A total of 93 completed responses were entered into the analysis. The results of the binary logit created coefficients for each attribute, which equated to the relative preference of the associated themes: clinical care, 2.022; education, 1.252; convenience, 1.245; ease of use, 1.155; communication, 1.040. All calculated coefficients were statistically significant ($P$<.01).

**Conclusions:** The results show that, in this cohort of patients with heart failure, the most preferred factor, clinical care, has enough value to be traded for approximately any two other factors. It also shows that the factor of communication is the least preferred attribute. Technology designers can use the associated preference weights to determine the relative increase of value perceived by patients by adding in certain attributes, with the greatest gains achieved by prioritizing clinical care. This would result in increased engagement in a chronic heart failure population that would benefit most from remote care.

***(JMIR Cardio 2025;9:e68022)*** doi:10.2196/68022

**KEYWORDS**

heart failure; telehealth; remote care; engagement; discrete choice; medical devices

## Introduction

Heart failure is a chronic and progressive condition, which is defined as the inability of the heart to pump sufficient blood to meet the body's oxygen demand. This is often caused by structural cardiac conditions that reduce the efficiency of the heart, for example, ischemic heart disease, because it weakens cardiac muscle and reduces the pump's effectiveness. Other associated conditions can contribute to the disease severity, such as diabetes mellitus and hypertension, which promote

structural changes to the heart, or chronic obstructive pulmonary disease, which reduces the blood's oxygen-carrying capacity. The result is a complex clinical syndrome that causes symptoms of fatigue, shortness of breath, and peripheral edema. As this usually occurs in an older patient cohort with an average age of 76 years with multiple comorbidities, their clinical management is complex and their health care needs are high. They typically have reduced mobility, cognition, and mood and face challenges in self-care and efficacy [1].

Remote care technologies can gather clinical data remotely, which enables closer monitoring of patients who are at a high risk of day-to-day clinical variation. These technologies provide easier access to care and have the potential to empower patients to improve self-management, enabling early identification and resolution of severe health issues before they require hospital admission. However, the drop-off rate for these devices is extremely high in this older population. Lack of engagement with the device may result in failure to achieve the anticipated improvements in clinical outcomes and could lead to a significant waste of time as well as research and development costs. This not only burdens patients and their health care providers but ultimately hinders the landscape of technology adoption in chronic diseases, limiting their potential to enhance patient care [2,3]. Therefore, when designing new remote care interventions, it is essential to consider user engagement as the driving force for the uptake and continued use of a remote care device for disease management.

A systematic review of the perceived benefits and drawbacks of remote care, from a clinician, patient, and carer viewpoint [4], identified five common themes that can be used to describe the experiences of users when engaging with remote care technology: communication (increasing interaction between patients and health care staff/carers/other patients), clinical care (improving the quality of care compared to established practice), education (providing tailored information to help with self-care and reduce uncertainty), ease of use (the technical aspects of the intervention are easy to handle without issues), and convenience (the intervention fits well around the patient's lifestyle and requires minimal effort). While this research concluded that each of the five themes was instrumental to maintaining patient engagement, it did not provide any insight as to which themes were prioritized most by patients. Therefore, to facilitate application of this work in real-world technology design, it is important to quantify the relative hierarchy of the themes and identify which factors could lead to greater engagement in a heart failure cohort.

Choice-based surveys can be used to understand the stated preference of a population for health provision [5]. Here we employ a discrete choice experiment (DCE) approach. In DCE, variables of interest or attributes are traded against one another in different scenarios to ascertain their relative importance [6]. These trade-offs provide information about patient decision-making processes in terms of what attributes participants are willing to compromise on in favor of others, thus understanding their ranked preference. DCEs can be used to simulate uptake or adoption of a new intervention or device based on its characteristics. This can also inform how changes in these attributes can affect user decisions under different scenarios and different values or levels of each attribute to determine to what extent changes should be made for optimal uptake.

We designed a DCE questionnaire to gather primary opinions from patients living with heart failure to elicit their preferences for remote care as categorized by our five themes. The themes capture user experience with minimal overlap and so are amenable to being delineated in questionnaire form, which lends itself well to a choice-based survey [7]. Using each theme as an attribute in the DCE design enables us to quantify the relative importance of each theme to patients with heart failure.

## Methods

### Overview

Since our themes were generated from grounded theory, their titles may be interpreted in a variety of ways. We therefore created clear descriptions for each attribute in relation to remote care (see Table 1). For each attribute, we chose two levels, positive and negative, corresponding to the level of attainment of any given attribute, with neutral included as a negative level [8,9].

**Table .** Description of the main attributes and levels used to determine the questions (trials) in the discrete choice questionnaire.[a]

| Attribute | Description | Level coding | Level description |
|---|---|---|---|
| Communication | The ability of the technology to create increased contact and follow up between patients and others, including health care staff, family, carers, or other patients | 0 | Reduces or does not improve opportunities for contact and communication |
| | | 1 | The technology increases opportunities for contacts and communication |
| Clinical care | The technology in some way affects the current clinical care given to the patient for their heart failure condition. | 0 | The technology makes no impact on current clinical care |
| | | 1 | Improves clinical care from current practice or provides more options for medical management, including providing information to make better decisions on care |
| Education | The impact of the technology on patients' knowledge about their health and self-care | 0 | There is no improvement in knowledge or ability to self-care |
| | | 1 | The technology provides details that clarify and provide useful information to the patient about their condition and aid in their self-care and management |
| Ease of use | The intuitiveness and relative ease that the technology can be introduced and used by new users, including technical difficulties and jargon | 0 | The technology is overly complex, with little technical support and may have a high rate of technical difficulties and complications, or is difficult to access for new users |
| | | 1 | The technology is easy and intuitive to use, requires relatively little support, or is easy to understand and use by a wide audience |
| Convenience | The measure of how much time and effort is saved by the use of the technology compared to normal care. Also relates to the level of comfort afforded by the technology in the patient's home. | 0 | There is no difference in the amount of time and effort required for self-care actions, or the device creates more work for the patient and requires extra time to use, or it creates increased worry or stress |
| | | 1 | The device functions to save time, such as automating processes or providing relevant information at the right time, and results in less work for self-care actions or allows the patient to be more comfortable in their own home environment |

[a]Attributes were taken from themes generated from a systematic thematic analysis of factors affecting user engagement with remote care technology in a population of patients with heart failure [4].

## Questionnaire Construction

Each question forced the participant to choose between two hypothetical remote care technologies with opposing attribute levels, that is, a positive level in an attribute in one choice means that the alternative choice will have a negative level of that same attribute. The forced choice design reduced the complexity of adding an opt-out alternative to each question, which minimized questionnaire fatigue [10].

The choice sets (the combination of levels of each attribute that were grouped together per question) were assigned based on a predetermined, orthogonal design algorithm [11]. For a discrete choice questionnaire containing 5 attributes each with 2 levels, this resulted in 32 profiles split across 16 questions. The order of the questions was randomized to mask the pattern of the choice sets. The attributes were listed in alphabetical order in each question [12,13].

## Sample Size

We used an established method for determining the minimum sample size for conjoint analyses [14]:

$$N > 500 \times ct \times a$$

Where *N* is the minimum sample size; *c* is the number of levels; *t* is the number of questions; and *a* is the number of alternative answers.

For a 16-question survey with 2 choices, the recommended minimum response size is 32 participants. We took this as a minimum and left the online survey open until the end of the study window to capture as many responses as possible.

### Criteria for Patient Participation

Patients who were aged 18 years or over and had a diagnosis of chronic heart failure were included in the study. Exclusion criteria included (a) diagnosis of acute heart failure without any chronic component and (b) non-English speaking patients (the questionnaire was only available in English).

### Patient and Public Involvement

A patient participation group consisting of five patients with heart failure and related conditions was formed to aid the outputs of the research. These patients were recruited via a free access public engagement event held at the University of Liverpool on October 23, 2017. This event involved talks from cardiology and technology experts to inform on upcoming heart failure technology research and generate interest in public participation. After the formation of the group, several informal discussions and feedback sessions were conducted between February and March 2018, where the group piloted the questionnaire and had input into the patient information leaflet. Design changes were made due to this feedback, including shortening the questions, formatting for better readability, as well as expanding on the patient information leaflet to provide more context for the study (Multimedia Appendix 1 and Multimedia Appendix 2). Furthermore, the patient group members helped to suggest places where the survey could be distributed online to heart failure care communities.

### Ethical Considerations

As per HRA guidance [15], responses to online surveys imply consent as long as participants are provided with sufficient information to reach an informed decision. We worked with our patient group to develop substantially descriptive participant information for them to make an informed choice. This study was approved by the Research Ethics Committee at the University of Liverpool (ref: 3314). The survey was exported online using a secure digital platform (SurveyMonkey), which complies with EU Privacy Laws and General Data Protection Regulations, and is registered under the Data Protection Act. This online platform was used to create a web link, which was the primary means of distributing the survey to participants. In accordance with the principles of data minimization and purpose limitation under General Data Protection Regulations, no personal or demographic data were collected by the research team; therefore, participants were not identifiable, nor was there any direct contact between the research staff and participants. No monetary compensation was offered to any participant for completing the questionnaire. Raw and processed data were stored securely on encrypted university intranet servers.

### Survey Distribution

The link to the survey was distributed to national and international heart failure patient groups, which were accessed via social media and communications through heart failure charities. A list of organizations approached for distribution can be found in the supplementary information (Multimedia Appendix 3). It is important to note that while the study was conducted in the United Kingdom, the survey was distributed worldwide, and so the respondents were not limited by geographic location.

### Analysis

Responses were analyzed using limited dependent-variable models to determine preference weights of each attribute [16]. From this, we can infer which attributes participants are willing to trade in favor of others. Our DCE is a forced-choice, five-attribute, two-level, two-alternative questionnaire (Multimedia Appendix 4). As both the choices and the levels were binary, binary logit [16] was used to determine the likelihood of the outcome. The logarithmic function ensures the likelihood values are constrained between 0 and 1 [17].

The logit definition is as follows: [18]:

$$\text{Logit}(P)=\log(\text{odds})=\log(P/(1-P))$$

As part of the regression, we assign logit(P) as a linear function of any given attribute *Xi*, so that:

$$\log P(1-P)=\alpha+\beta Xi=Ui$$

Where *P*=probability (of choosing this option); α=reference value or constant; β=coefficient of attribute *X*; *i*=attribute number; *U*=utility

The logit value is proportional to the odds of an attribute, affecting the probability of choosing an alternative. Thus, these values can be compared directly as preference weights for each variable. The preference value for each attribute is known as utility, which is the measure of importance of each attribute or combination of attributes. In order to standardize for participant heterogeneity, random effects were added to create a mixed binary logit model [18,19].

The utility value of each combination of attribute level was obtained by adding the constant coefficient of attribute *X* from the logit model, with the coefficients of each positive attribute present. The odds were obtained by exponentiating the utility. To convert this to percentage uptake probability, that is, the likelihood of choosing this remote care device as opposed to the alternative, we divided the Odds by 1+Odds [20]. The dataset was analyzed using RStudio version 1.0.136. These calculations were also corroborated using STATA/MP 13.0.

## Results

The survey was open for 133 days (June 3, 2018–October 14, 2018) and was initiated by 164 participants. The completion rate was 57%, giving 94 completed responses. A limited trial of the paper questionnaire was undertaken in local heart failure clinics, but this generated only 1 completed response. To verify accuracy and consistency of the extracted results, visual inspection was undertaken to assess for discrepancies and

anomalous data, and all survey attempts with missing or incomplete responses were excluded. Response nondifferentiation was identified, and two responses were omitted due to nontrading (all responses from a participant were either choice A or choice B). This left 93 valid responses. (Multimedia Appendix 5)

We identified some positive attribute dominance in the responses (respondent always chose the option with a positive level in a single attribute) [21]: 10 participants had positive dominance for clinical care, three for education, two for ease of use, and one for communication. There were no cases of negative attribute dominance. The main outputs of the mixed binary logit are displayed in Table 2.

**Table .** Results from the binary logit analysis of the discrete choice questionnaire[a].

| Attribute | Coefficient (95% CI) | *P* value |
|---|---|---|
| *Intercept* | −3.357 (−3.654 to −3.060) | <.001 |
| Clinical care | 2.022 (1.810 to 2.233) | <.001 |
| Education | 1.252 (1.077 to 1.428) | <.001 |
| Convenience | 1.245 (1.053 to 1.436) | <.001 |
| Ease of use | 1.155 (0.982 to 1.327) | <.001 |
| Communication | 1.040 (0.864 to 1.216) | <.001 |

[a]The coefficients for each attribute represent relative patient preference weighting for that attribute in isolation, relative to the intercept. Higher value coefficients represent a proportional increase in preference by patients with heart failure.

Each coefficient was highly statistically significant, indicating that there was a sufficient sample size and significant effect of each attribute on patient choice. The goodness of fit was evaluated using the pseudo R-squared of the logit model, which showed a value of 0.1833. The attributes presented in the model thus explain 18% of the variance in choice of each participant, a typical result for a DCE of this size [22].

We calculated the utility value, odds ratio, and percentage probability of choosing each combination of attribute levels (Table 3). The utility represents the preference value for choosing each alternative and can be compared for evaluating complete choice sets (different combinations of attributes). This contrasts with coefficient values for each attribute, calculated from the logit model, which indicates preferences for individual attributes (Textbox 1).

**Table .** A comparison of all 32 possible combinations of attributes and levels that can be applied to a remote care intervention.

| Communication | Clinical care | Education | Ease of use | Convenience | Utility | Odds | % uptake probability |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 3.36 | 28.70 | 96.63 |
| 0 | 1 | 1 | 1 | 1 | 2.32 | 10.14 | 91.02 |
| 1 | 1 | 1 | 0 | 1 | 2.20 | 9.05 | 90.05 |
| 1 | 1 | 1 | 1 | 0 | 2.11 | 8.27 | 89.21 |
| 1 | 1 | 0 | 1 | 1 | 2.10 | 8.20 | 89.13 |
| 1 | 0 | 1 | 1 | 1 | 1.34 | 3.80 | 79.17 |
| 0 | 1 | 1 | 0 | 1 | 1.16 | 3.20 | 76.17 |
| 0 | 1 | 1 | 1 | 0 | 1.07 | 2.92 | 74.49 |
| 0 | 1 | 0 | 1 | 1 | 1.06 | 2.90 | 74.35 |
| 1 | 1 | 1 | 0 | 0 | 0.96 | 2.61 | 72.26 |
| 1 | 1 | 0 | 0 | 1 | 0.95 | 2.59 | 72.11 |
| 1 | 1 | 0 | 1 | 0 | 0.86 | 2.36 | 70.26 |
| 0 | 0 | 1 | 1 | 1 | 0.29 | 1.34 | 57.32 |
| 1 | 0 | 1 | 0 | 1 | 0.18 | 1.20 | 54.50 |
| 1 | 0 | 1 | 1 | 0 | 0.09 | 1.09 | 52.26 |
| 1 | 0 | 0 | 1 | 1 | 0.08 | 1.09 | 52.07 |
| 0 | 1 | 1 | 0 | 0 | −0.08 | 0.92 | 47.93 |
| 0 | 1 | 0 | 0 | 1 | −0.09 | 0.91 | 47.74 |
| 0 | 1 | 0 | 1 | 0 | −0.18 | 0.83 | 45.50 |
| 1 | 1 | 0 | 0 | 0 | −0.29 | 0.74 | 42.68 |
| 0 | 0 | 1 | 0 | 1 | −0.86 | 0.42 | 29.74 |
| 0 | 0 | 1 | 1 | 0 | −0.95 | 0.39 | 27.89 |
| 0 | 0 | 0 | 1 | 1 | −0.96 | 0.38 | 27.74 |
| 1 | 0 | 1 | 0 | 0 | −1.06 | 0.35 | 25.65 |
| 1 | 0 | 0 | 0 | 1 | −1.07 | 0.34 | 25.51 |
| 1 | 0 | 0 | 1 | 0 | −1.16 | 0.31 | 23.83 |
| 0 | 1 | 0 | 0 | 0 | −1.34 | 0.26 | 20.83 |
| 0 | 0 | 1 | 0 | 0 | −2.10 | 0.12 | 10.87 |
| 0 | 0 | 0 | 0 | 1 | −2.11 | 0.12 | 10.79 |
| 0 | 0 | 0 | 1 | 0 | −2.20 | 0.11 | 9.95 |
| 1 | 0 | 0 | 0 | 0 | −2.32 | 0.10 | 8.98 |
| 0 | 0 | 0 | 0 | 0 | −3.36 | 0.03 | 3.37 |

[a]The table compares each intervention's relative utility, odds ratio and percentage uptake probability values, which can each be considered as composite preference weights of the combination of all attribute levels in a remote care intervention.

**Textbox 1.** How to use the data for comparative analysis as a worked example.

The percentage uptake probabilities are derived from the calculated utility score and so are symmetrical, giving a probability of 50% to an intervention with a utility score of 0. As such, they are not intended to be used in isolation but mainly as a means of calculating marginal differences in engagement between two comparator intervention states.

To compare engagement between two different types of intervention, for example, with and without a certain attribute included, we can use Table 3 to calculate the marginal probability, which is the difference in percentage uptake probability between the two interventions. This can be done by choosing the two rows that most correspond to each individual remote care device (based on present attributes) and then subtracting the percentage uptake probabilities from each other to get the marginal probability.

For example, in a remote care intervention with no attributes present (row: 0/0/0/0/0), the percentage uptake probability is 3.37%. An intervention that has the attribute of communication alone (row: 1/0/0/0/0) has the percentage uptake probability of 8.98%. Therefore, the marginal probability gained by adding the communication attribute to the intervention which has no attributes is calculated as 8.98−3.37=+5.61%.

Alternately, to work out the marginal probability of adding clinical care instead, we look to the row which only includes the clinical care attribute (row: 0/1/0/0/0) to see that its percentage uptake probability is 20.83%. We then subtract this from the percentage uptake probability of the intervention with no attributes: 20.83–3.37=+17.46%.

The marginal probability figure can be regarded as the change in utility between comparator interventions and represents the amount of value added in terms of engagement by altering the remote care intervention to meet specific additional attributes. At a glance, it can therefore be seen that the value added from incorporating the clinical care attribute is much greater than adding the communication attribute to an intervention without either.

Taking the mean of marginal probabilities for adding the attribute to each permutation which excludes it gives another quantitative measure of patient preference. We found the mean marginal probabilities per attribute to be as follows: communication=+18.04%, ease of use=+20.1%, convenience=+21.8%, education=+21.9%, and clinical care=+37.6%. These values could also be interpreted as the average relative increase in patient preference gained by adding this attribute to an intervention that lacks it. This is a useful measure for comparing the value of the attributes themselves against each other; however, for a more detailed comparison of combinations of attributes (whole interventions), the marginal probability described in the above calculation would be more suitable. For example, mean marginal probabilities suggest patients would be more likely to value adding clinical care outcomes to an intervention (+37.6%) compared to adding communication to an intervention (+18.0%) on average. However, if the aim is to compare an intervention with no attributes and one which has both clinical care and communication, the specific marginal probability between these interventions can be calculated more precisely. Refer to the row that contains both clinical care and communication (row: 1/1/0/0/0) to see that the percentage uptake probability for this intervention is 42.68%. Then calculate the difference between this and the percentage uptake probability of the intervention with no attributes as in the examples above (row: 0/0/0/0/0). This gives a marginal probability of 42.68−3.37=+39.31%. Thus, the specific marginal probabilities are ideal to be used when there is a fixed intervention state, or a starting point, such as a design or existing intervention that is intended to be improved upon.

# Discussion

## Principal Findings

The analysis ranked the remote care attributes in the following order of importance for patients with heart failure: (1) clinical care; (2) education; (3) convenience; (4) ease of use; and (5) communication. Based on the coefficients of the logit fit, clinical care was almost twice as important as the lowest scoring variable, communication. Remote care technology design should therefore prioritize clinical care improvements first and foremost. The attributes of education and convenience had similar preference values, which were around 20% greater than communication. Ease of use was 11% more important than communication. This pattern of preference shows a disproportionately high preponderance toward clinical care, with the second, third, and fourth ranked attributes plateauing at a similar level. Therefore, if a trade-off is required, any other attribute may be sacrificed for the sake of preserving clinical care, while still incentivizing patient engagement.

## Comparison to Prior Work

A number of DCE and conjoint analyses have been published regarding patient preferences for telecare since the COVID pandemic [23-25]. However, there have been no other DCEs evaluating engagement of remote care technologies in this patient cohort of chronic heart failure. Therefore, the study provides a valuable insight into the factors of remote care devices that encourage engagement. In a post-COVID era, remote care technologies have gained greater importance in health care. Patients with heart failure are a vulnerable cohort and so are more likely to be offered remote consultation. Therefore, these preference rankings are all the more vital at this time to help remote care become better established in medical practice for those that need it most.

## Strengths and Limitations

### *Methodological Design Advantages*

Among the advantages of our experiment was that each possible combination of levels and attributes was presented to the participants, resulting in a full factorial design. This establishes a more accurate statistical value for each preference as fewer assumptions are made. By contrast, partial factorial designs sacrifice comprehensiveness for brevity [26].

Another strength is that the attributes used were based on evidence from a grounded theory qualitative systematic review, specific to the subject [4]. This means that the outputs of the review were tailored to this questionnaire design, resulting in relevant attributes derived from high-quality evidence.

### *Questionnaire Considerations*

Our study does have some limitations. First, the statistical model assumes each participant will always choose the option that maximizes their utility, which could lead to bias. We tried to mitigate this by adding a random effect to model heterogeneity of preference choices, even if they might be irrational (or of less utility). This study, therefore, presents the preference values

in terms of a probability of choosing each option, which means the likelihood of a nonrational choice still exists.

Second, the DCE assumes that the participant is equally attentive on question 1 as they are on question 16, and this may not always be the case. The complexity of the questions coupled with their repetitive nature may contribute to participant fatigue when answering questions. We had the option of creating either an 8-question design or a 16-question design. We opted for the latter to obtain a greater statistical effect from each respondent. In hindsight, this may have contributed to the high non-completer rates [27].

Third, in many DCEs, the alternative choices are based on existing interventions or ones that are ready to market. In this study, we asked participants to imagine theoretical technologies. This enables the outputs to be applied to a wide variety of technology designs in the future. A disadvantage is the potential for hypothetical bias, which can lead to a discrepancy between patient stated preference and the actual (or revealed) preference [28].

Fourth, related to this hypothetical scenario is the fact that an opt-out option was not presented to participants. This forced-choice design meant that they were not able to express dissatisfaction with both alternatives at once. We recognize that this is an artificial scenario, and in reality, participants may be disinclined to engage with either option. However, the aim of this study was to understand the ranking preferences of patient behavior rather than whether they would engage in any specific intervention. Thus, the design of the study was adapted to maximize the depth of information, at the cost of using hypothetical scenarios.

Finally, there was a lack of a third *neutral* level for each attribute: either the attribute was present in the remote care technology or it was not. This means that there was no *neither* option for the participant to choose to indicate that a specific attribute was unimportant in their decision-making. Furthermore, the negative level was often used to effectively indicate two different levels by specifying both an *absence* and *negative* effect of the attribute within the meaning. Although we chose to omit the neutral level from the questionnaire design, the benefit of this is that it allows the analysis to be more straightforward in terms of the binary logit analysis rather than adopting a multinomial logit model, which requires more assumptions [29]. Another benefit to the two-level system was that the choice burden on the participants was minimized which likely improved completion rates.

### *Generalizability*

First, the effects of the recorded attributes are presented in relation to one another, which means that the assessments of value lack generalizability outside of the context of the comparison versus each other in a heart failure cohort. It is important, therefore, to realize that these results may not translate to cohorts with other conditions, or even other chronic diseases, and that the results do not have intrinsic value independent from the attributes they are compared to here. A mitigating factor is that the analysis relies on the foundation of its supporting research to substantiate the list of tested attributes. The supporting research is a thorough and in-depth look at lived experiences within this cohort and seeks to be as comprehensive as possible while capturing commonalities in themes that can be of value in the assessment of technology in this chronic condition [6].

Second, the online self-selection method may reduce the generalizability of the study findings to other cohorts such as in-person heart failure clinics. It was likely also completed by those with greater digital access and skills. However, in a post-COVID era, where patients are more likely to be familiar with remote care, those lacking digital access and skills may be in the minority. Our findings should nevertheless be interpreted within the context of patients who are generally supportive of new technologies [30].

Third, the methodology used in this study resulted in a lack of demographic data collection. This may also deter from the generalizability of findings. While the heart failure demographic is generally well established, the self-selection and timing of the questionnaire, as well as its online distribution route, have the potential to skew the responses based on whether the participant sample was seen to diverge from the average heart failure demographic, for example, to those younger, with less comorbidities, living in more affluent locations. Without the demographic data to put these results into context, the generalizability of the outputs when applied to a new cohort of patients with heart failure may be affected. However, since the attributes were built from a rigorous analysis of patient experience data generated from a variety of patient demographics and geographical locations, we posit that the central themes continue to have relevance across a wide range of patient populations.

Finally, the factor of *cost*, which is normally assessed in this manner by means of adding an attribute that asks how much the participant is willing to pay for certain factors, is missing. The remote care intervention that participants were asked to envision was hypothetical, and therefore there is no real-world cost to incorporate in the assessment. The same can be said for other real-world factors such as management, administration, and access to the intervention. This may potentially lead to inaccurate responses as the hypothetical scenarios may pose unrealistic cost choices with reduced credibility effect, leading to invalid willingness to pay estimates [31]. However, it is worthwhile considering that cost implications and access restrictions played a role in defining the attribute of *ease of use* in the original thematic synthesis, as high cost and maintenance requirements of the device contributed to poor accessibility of the intervention and was seen to impact the ease of use for patients [32].

## Future Work

Improving this and similar surveys may require shifts in methodologies to make it more generalizable, albeit with additional feedback. In the first instance, while patient participation was a key determinant in the design of the methodology, additional value may have been obtained by reaching out to technology designers and start-ups that create devices within the space. Obtaining this kind of feedback would enable tailoring of the outputs in such a way as to provide the

most deliverable benefit in the context of future design by, for instance, presenting realistic alternatives grounded in existing technologies as opposed to theoretical ones.

To mitigate some of the limitations further, it may be also useful to obtain demographic information and location of participants so as to correctly contextualize the responses based on patient profile, recognizing that different subpopulations may have differences in preference.

Finally, in order to address noncompletion rates, the questionnaire could be shortened in order to be less mentally taxing, while also ensuring a process of gathering feedback from participants as to reasons for noncompletion.

## Conclusions

Our questionnaire used a DCE method to elicit preferences for remote care technology from patients with heart failure from around the world. Results of the analysis indicate that clinical care was substantially more valued as a factor for engagement with remote technology than the four other themes of education, convenience, ease of use, and communication. Our findings also allow approximations of increase in engagement by sequentially adding in these individual factors to an existing remote care device based on their preference values. This hierarchy could provide useful insights for technology designers to check the effectiveness of an intervention's features in engaging the end user and help develop a plan of improvement for devices based on their missing attributes. Incorporating these attributes appropriately will ultimately bring remote care technology to these patients in a more effective and engaging manner, to reduce the burden of morbidity from chronic heart failure.

## Acknowledgments

We acknowledge the vital input from the heart failure patient group at Liverpool Heart and Chest Hospital in overseeing the questionnaire and patient information leaflets. We would particularly like to thank Lynn Hedgecoe, who was instrumental in distributing the survey to patient groups as well as reviewing the manuscript from a patient perspective. Generative AI was not used in any part of this work. This work was supported by the National Institute for Health Research (NIHR), Applied Research Collaboration, North West Coast. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care. For the purpose of Open Access, the author has applied a Creative Commons Attribution (CC-BY) license to any Author Accepted Manuscript version arising.

## Data Availability

The raw data output is available as supplementary information (Multimedia Appendix 5).

## Authors' Contributions

AAN was involved in every stage of the project. JD, DH, and MP contributed to conceptualization, funding acquisition, supervision, and review and editing, with DH additionally contributing to methodology, formal analysis, and validation.

## Conflicts of Interest

MP currently receives partnership funding, paid to the University of Liverpool, for the following: Medical Research Council (MRC) Clinical Pharmacology Training Scheme (co-funded by MRC and Roche, UCB, Eli Lilly, and Novartis) and the MRC Medicines Development Fellowship Scheme (co-funded by MRC and GSK, AstraZeneca, Optum, and Hammersmith Medicines Research). He has developed a genotyping panel with MC Diagnostics but does not benefit financially from this. He is part of the Innovative Medicines Initiative Consortium: Accelerating Research & Development for Advanced Therapies [33]; none of these funding sources have been used for the current research. AAN is currently employed by Novo Nordisk; however, this research was carried out in full prior to this appointment, and at the time of the study, AA had no affiliations with Novo Nordisk or other commercial interests.

Multimedia Appendix 1
Participant information sheet for discrete choice experiment.
[DOCX File, 792 KB - cardio_v9i1e68022_app1.docx ]

Multimedia Appendix 2
Instructions for answering questionnaire.
[DOCX File, 120 KB - cardio_v9i1e68022_app2.docx ]

Multimedia Appendix 3
Charities, organizations, and social groups contacted for online questionnaire distribution.
[DOCX File, 17 KB - cardio_v9i1e68022_app3.docx ]

Multimedia Appendix 4
Discrete choice experiment questionnaire.
[DOCX File, 46 KB - cardio_v9i1e68022_app4.docx ]

Multimedia Appendix 5
Online questionnaire responses.
[DOCX File, 44 KB - cardio_v9i1e68022_app5.docx ]

## References


1. Al-Naher A, Wright D, Devonald MAJ, Pirmohamed M. Renal function monitoring in heart failure - what is the optimal frequency? A narrative review. Br J Clin Pharmacol 2018 Jan;84(1):5-17. [doi: 10.1111/bcp.13434] [Medline: 28901643]
2. Thomas EE, Taylor ML, Banbury A, et al. Factors influencing the effectiveness of remote patient monitoring interventions: a realist review. BMJ Open 2021 Aug 25;11(8):e051844. [doi: 10.1136/bmjopen-2021-051844] [Medline: 34433611]
3. McLean S, Sheikh A, Cresswell K, et al. The impact of telehealthcare on the quality and safety of care: a systematic overview. PLOS ONE 2013;8(8):e71238. [doi: 10.1371/journal.pone.0071238] [Medline: 23977001]
4. Al-Naher A, Downing J, Scott KA, Pirmohamed M. Factors affecting patient and physician engagement in remote health care for heart failure: systematic review. JMIR Cardio 2022 Apr 6;6(1):e33366. [doi: 10.2196/33366] [Medline: 35384851]
5. Ryan M, Watson V. Comparing welfare estimates from payment card contingent valuation and discrete choice experiments. Health Econ 2009 Apr;18(4):389-401. [doi: 10.1002/hec.1364] [Medline: 18677721]
6. Ryan M, Bate A, Eastmond CJ, Ludbrook A. Use of discrete choice experiments to elicit preferences. Qual Health Care 2001 Sep;10 Suppl 1(Suppl 1):i55-i60. [doi: 10.1136/qhc.0100055] [Medline: 11533440]
7. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
8. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health Econ 2012 Jun;21(6):730-741. [doi: 10.1002/hec.1739] [Medline: 21557381]
9. Penn J. The effect of forced choice with constant choice experiment complexity. Presented at: Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting; Jul 27-29, 2014; Minneapolis URL: http://ageconsearch.umn.edu/bitstream/169777/2/2014%20AAEA%20paper.pdf [accessed 2025-10-31]
10. Le A, Han BH, Palamar JJ. When national drug surveys "take too long": an examination of who is at risk for survey fatigue. Drug Alcohol Depend 2021 Aug 1;225:108769. [doi: 10.1016/j.drugalcdep.2021.108769] [Medline: 34049103]
11. Kocur G, Adler T, Hyman W, Aunet B. Guide to forecasting travel demand with direct utility assessment. : Administration UMT; 1982 URL: https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/PB82200270.xhtml [accessed 2025-10-31]
12. Lancsar E, Fiebig DG, Hole AR. Discrete choice experiments: a guide to model specification, estimation and software. Pharmacoeconomics 2017 Jul;35(7):697-716. [doi: 10.1007/s40273-017-0506-4] [Medline: 28374325]
13. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health 2013;16(1):3-13. [doi: 10.1016/j.jval.2012.08.2223] [Medline: 23337210]
14. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015 Oct;8(5):373-384. [doi: 10.1007/s40271-015-0118-z] [Medline: 25726010]
15. Applying a proportionate approach to the process of seeking consent. : HRA; 2017 URL: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/applying-proportionate-approach-process-seeking-consent_R3gbJKn.pd [accessed 2025-10-31]
16. Hauber AB, González JM, Groothuis-Oudshoorn CGM, et al. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR conjoint analysis good research practices task force. Value Health 2016 Jun;19(4):300-315. [doi: 10.1016/j.jval.2016.04.004] [Medline: 27325321]
17. Eby WM TM. Robust logistic and probit methods for binary and multinomial regression. J Biom Biostat 2014;05(4). [doi: 10.4172/2155-6180.1000202]
18. Berger D. Introduction to Binary Logistic Regression and Propensity Score Analysis: Claremont Graduate University; 2017. URL: https://wise.cgu.edu/wp-content/uploads/2016/07/Introduction-to-Logistic-Regression.pdf [accessed 2025-10-31]
19. Parzen M, Ghosh S, Lipsitz S, et al. A generalized linear mixed model for longitudinal binary data with a marginal logit link function. Ann Appl Stat 2011;5(1):449-467. [doi: 10.1214/10-AOAS390] [Medline: 21532998]
20. Sperandei S. Understanding logistic regression analysis. Biochem Med (Zagreb) 2014;24(1):12-18. [doi: 10.11613/BM.2014.003] [Medline: 24627710]
21. Lindberg MA. The development of attribute dominance in the knowledge base. J Genet Psychol 1989 Sep;150(3):269-280. [doi: 10.1080/00221325.1989.9914596] [Medline: 2809574]
22. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete choice experiments in health economics: past, present and future. Pharmacoeconomics 2019 Feb;37(2):201-226. [doi: 10.1007/s40273-018-0734-2] [Medline: 30392040]

23. Mozes I, Mossinson D, Schilder H, Dvir D, Baron-Epel O, Heymann A. Patients' preferences for telemedicine versus in-clinic consultation in primary care during the COVID-19 pandemic. BMC Prim Care 2022 Feb 22;23(1):33. [doi: 10.1186/s12875-022-01640-y] [Medline: 35193509]
24. Tierney AA, Brown TT, Aguilera A, Shortell SM, Rodriguez HP. Conjoint analysis of telemedicine preferences for hypertension management among adult patients. Telemed J E Health 2024 Mar;30(3):692-704. [doi: 10.1089/tmj.2023.0254] [Medline: 37843962]
25. Gabay G, Ornoy H, Moskowitz H. Patient-centered care in telemedicine-an experimental-design study. Int J Med Inform 2022 Mar;159:104672. [doi: 10.1016/j.ijmedinf.2021.104672] [Medline: 34979434]
26. Dobra RA, Boeri M, Elborn S, Kee F, Madge S, Davies JC. Discrete choice experiment (DCE) to quantify the influence of trial features on the decision to participate in cystic fibrosis (CF) clinical trials. BMJ Open 2021 Mar 2;11(3):e045803. [doi: 10.1136/bmjopen-2020-045803] [Medline: 33653764]
27. Bansback N, Hole AR, Mulhern B, Tsuchiya A. Testing a discrete choice experiment including duration to value health states for large descriptive systems: addressing design and sampling issues. Soc Sci Med 2014 Aug;114(100):38-48. [doi: 10.1016/j.socscimed.2014.05.026] [Medline: 24908173]
28. Liljas B, Blumenschein K. On hypothetical bias and calibration in cost-benefit studies. Health Policy 2000 May;52(1):53-70. [doi: 10.1016/s0168-8510(00)00067-1] [Medline: 10899644]
29. Mangham LJ, Hanson K, McPake B. How to do (or not to do)... Designing a discrete choice experiment for application in a low-income country. Health Policy Plan 2009 Mar;24(2):151-158. [doi: 10.1093/heapol/czn047] [Medline: 19112071]
30. Cravo Oliveira T, Barlow J, Bayer S. The association between general practitioner participation in joint teleconsultations and rates of referral: a discrete choice experiment. BMC Fam Pract 2015 Apr 21;16:50. [doi: 10.1186/s12875-015-0261-6] [Medline: 25896515]
31. Genie MG, Ryan M, Krucien N. To pay or not to pay? Cost information processing in the valuation of publicly funded healthcare. Soc Sci Med 2021 May;276:113822. [doi: 10.1016/j.socscimed.2021.113822] [Medline: 33752103]
32. Rowen D, Stevens K, Labeit A, et al. Using a discrete-choice experiment involving cost to value a classification system measuring the quality-of-life impact of self-management for diabetes. Value Health 2018 Jan;21(1):69-77. [doi: 10.1016/j.jval.2017.06.016] [Medline: 29304943]
33. Ardat. URL: www.ardat.org [accessed 2025-10-31]


## Abbreviations

**DCE:** Discrete choice experiment
**MRC:** Medical Research Council
**NIHR:** National Institute for Health Research


*Edited by A Coristine; submitted 26.10.24; peer-reviewed by B Tyl, B Keen; revised version received 26.07.25; accepted 30.07.25; published 05.11.25.*

*Please cite as:*
*Al-Naher A, Downing J, Hughes D, Pirmohamed M*
*Patient Preferences for Using Remote Care Technology in Heart Failure: Discrete Choice Experiment*
*JMIR Cardio 2025;9:e68022*
*URL: https://cardio.jmir.org/2025/1/e68022*
*doi:10.2196/68022*

# Prerequisites for Cost-Effective Home Blood Pressure Telemonitoring: Early Health Economic Analysis

Job van Steenkiste[1,2,3*], MD; Pim van Dorst[4,5*], MSc; Daan Dohmen[3], Prof Dr; Cornelis Boersma[4,5], Prof Dr

[1]Department of Internal Medicine, Maasstad Hospital, Rotterdam, The Netherlands

[2]Department of Hospital Pharmacy, Erasmus MC University Medical Center, Rotterdam, The Netherlands

[3]Faculty of Management Sciences, Open University, Valkenburgerweg 177, Heerlen, The Netherlands

[4]Department of Health Sciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[5]Health-Ecore Ltd, Zeist, The Netherlands

[*]these authors contributed equally

**Corresponding Author:**
Job van Steenkiste, MD
Department of Internal Medicine, Maasstad Hospital, Rotterdam, The Netherlands

## Abstract

**Background:** Home blood pressure telemonitoring (HBPT) has been proposed to enhance adherence and optimize health care delivery, yet its prerequisites for cost-effective implementation remain unclear.

**Objective:** This study aims to quantify the potential cost-effectiveness of HBPT and identify prerequisites for cost-effective implementation of HBPT in comparison to standard hypertension management, using an early health economic analysis from a societal perspective.

**Methods:** A decision-analytic Markov model with a lifetime horizon (30 years) and a willingness-to-pay threshold of €20,000 (€1=US $1.09) per quality-adjusted life year (QALY) was developed to assess the cost-effectiveness of HBPT compared to standard of care (SOC). The HBPT intervention was based on an existing HBPT program applied by the Maasstad Hospital, Rotterdam, the Netherlands. The model incorporated 12 health states: 7 blood pressure states, 1 cardiovascular (CV) event, 1 recurrent CV event, 1 postrecurrent CV event, 1 all-cause death, and 1 CV disease–related death. A hypothetical cohort of 1000 patients (average age 65.3 years) was modeled, and results were reported in costs, QALYs, and the incremental cost-effectiveness ratio (ICER). The model assumed 3 in-person outpatient department (OPD) consultations in the SOC group and 1.5 in the HBPT group. Extensive sensitivity analyses were performed to identify important variables for the cost-effective implementation of HBPT.

**Results:** Following the base-case analysis, HBPT was not cost-effective with an ICER of €20,386 per QALY. Sensitivity analyses indicated that reducing the number of in-person OPD consultations resulted in a more favorable ICER. Specifically, reducing the number of in-person OPD consultations to 1.48 annually resulted in an ICER below the willingness-to-pay threshold. Reducing the in-person OPD consultations to an average of 1.18 per year would make HBPT cost-saving. Scenario analyses revealed that extending the duration of HBPT's clinical effect to 2 or 3 years substantially improved the ICER. Additionally, targeting HBPT toward patients aged 64 years or below further improved the ICER.

**Conclusions:** HBPT could result in cost-effective or cost-saving outcomes with only minor reductions in in-person OPD consultations. These findings highlight the potential of HBPT to transform hypertension management by replacing traditional hypertension management with more efficient care using remote patient monitoring.

***(JMIR Cardio 2025;9:e64386)*** doi:10.2196/64386

**KEYWORDS**

hypertension; blood pressure; telemonitoring; cost-effectiveness; economic evaluation; monitoring; health economics; cost; cost-effective; management; cardiovascular disease; intervention; lifestyle; adherence; clinical trials

## Introduction

Hypertension remains one of the most important risk factors for cardiovascular diseases (CVDs) [1]. Despite lifestyle and drug therapy interventions, a significant proportion of patients with hypertension remains inadequately controlled, which is mostly the result of poor medication adherence [2]. Home blood pressure telemonitoring (HBPT) has been proposed to improve adherence [2,3] by allowing patients to measure their blood pressure at home while being remotely monitored by their health care providers. Proactive monitoring in patients with off-target blood pressures could improve overall blood pressure control

through adjustment of medical treatments or by improving adherence, in particular to drug therapy [4]. Besides its potential to improve clinical outcomes, HBPT could optimize health care delivery and resource use [5] by including patient-specific measurement schedules and monitoring algorithms, designed by the responsible health care providers. Automated alerts could inform the clinician if the patient remains off-target, thereby drawing the clinician's attention to those patients who need it the most. Furthermore, modern-day telemonitoring platforms (eg, Luscii [6] and Patient Journey App [7]) do not solely provide measuring and monitoring functionalities but also serve as a platform for digital coaching and education on lifestyle factors that can further improve clinical outcomes [3].

Recent clinical evidence on HBPT confirms positive effects on blood pressure control [8], but widespread adoption of HBPT is still limited in the Netherlands. One of the perceived barriers to large-scale implementation is the lack of a clear reimbursement structure, which is related to a lack of evidence on the cost-effectiveness of this digital health intervention [9]. Clinical trials evaluating the effectiveness of HBPT often have limited follow-up durations [8]. Consequently, they only demonstrate short-term benefits on blood pressure control and do not capture the potential long-term advantages, such as reductions in cardiovascular (CV) events. Furthermore, available evaluations of HBPT in patients with hypertension mainly focus on the cost impact of HBPT, do not report on the impact of HBPT on the quality of life of the patient [10,11], and are not representative of the Dutch hospital setting [12]. Hence, there is a need to quantify the long-term value of HBPT in terms of costs and health outcomes while considering the limited data availability on resource use and effectiveness.

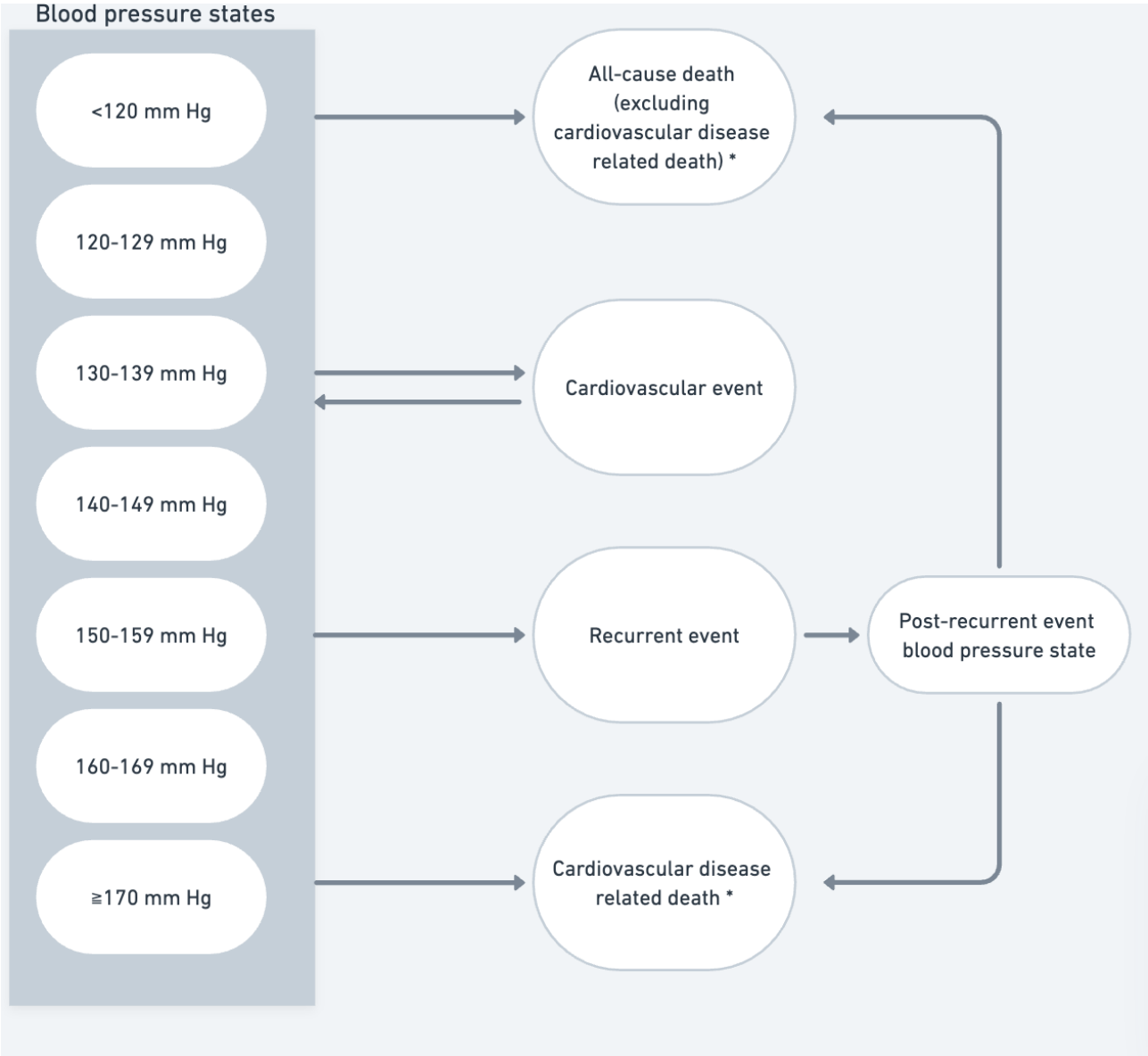
In this study, we aim to quantify the potential of HBPT in terms of cost-effectiveness with an early health economic analysis in patients with hypertension. Additionally, we aim to identify important prerequisites for cost-effective implementation of HBPT.

## Methods

### Study Design

This early health economic evaluation is reported per the 2022 CHEERS (Consolidated Health Economic Evaluation Reporting Standards) guidelines for reporting economic evaluations (Checklist 1). Given the lack of long-term efficacy data of HBPT [8] and the resulting uncertainty in the clinical evidence, the current evaluation is considered an early health economic evaluation, which is based on available literature [13]. A decision-analytic Markov model (see Figure 1) with a lifetime (30 years) horizon and a willingness-to-pay (WTP) threshold of €20,000 (€1=USD $1.09) per quality-adjusted life year (QALY) [14], to assess the cost-effectiveness of HBPT in combination with drug therapy for patients with hypertension. A societal perspective was applied (eg, including direct medical costs and nonmedical costs) according to the Dutch guideline for conducting health economic research [15]. All costs were inflated using Dutch inflation rates to reflect the costs in 2024 euro [16]. The model was developed in R statistical software version 4.4.1 [17].

**Figure 1.** Markov model including 12 different health states. *Risk of all-cause death and cardiovascular disease–related death were blood pressure independent.




## Model Overview

### *Structure*

The model included a hypothetical population of 1000 patients with an average age of 65.3 years and consisted of 12 health states: <120, 120 - 129, 130 - 139, 140 - 149, 150 - 159, 160 - 169, ≥170 mm Hg, CV event, recurrent event, postrecurrent event, all-cause death, and CVD-related death. The initial distribution and demographics (Multimedia Appendix 1 [8,18-23]) of patients over the systolic blood pressure states were based on the Blood Pressure Lowering Treatment Trialists' Collaboration study [18], reflecting a real-world distribution of patients with hypertension and no history of CVD. Patients could transition from higher blood-pressure states to lower blood-pressure states on an annual basis, based on drug therapy, until the patients were on target (120 - 129 mm Hg). A reverse transition was not possible. Each year, patients could experience a CV event, after which the patients returned to the pre-event blood-pressure health state, as no direct blood pressure lowering effect due to the CV event was expected. A CV event was a composite event consisting of either a myocardial infarction (MI), a cerebral hemorrhage, or an ischemic cerebrovascular event, and the event risk was blood pressure dependent [19]. Patients could experience a recurrent event after which they progressed into the postrecurrent event health state. The risks of all-cause death and CVD-related death were assumed to be blood pressure independent.

### *Standard of Care*

Standard of care (SOC) was based on the current practice of hypertension management in the Netherlands. The care provided via the hospital outpatient department (OPD) was based on the latest European Society of Hypertension guidelines on the management of arterial hypertension [24]. In the model, patients in the SOC group would be managed with drug therapy and lifestyle interventions. Patients would on average have 3 in-person OPD consultations in the hospital with their clinician during each 1-year cycle, based on the standard

diagnosis-treatment combination for patients with hypertension in the Netherlands [25].

### *Intervention*

Patients in the HBPT group were similarly managed in terms of drug treatment compared to the SOC group. The HBPT intervention was based on the HBPT program developed by the Maasstad Hospital in Rotterdam, the Netherlands, and adopted and studied throughout the region [26]. In this program, which is conducted in a hospital setting, patients measured their blood pressure during a complete week with 2 measurements in the morning and 2 in the evening. Measurement weeks were scheduled depending on the level of blood pressure control (eg, weekly in case of very uncontrolled blood pressure [>180/110 mm Hg] and monthly in case of controlled blood pressure [<140/90 mm Hg]), but on average occurred once every month. The monitoring platform used was the Luscii [6] application, which is the most widely used platform in the Netherlands for remote patient monitoring. The most frequently used patient monitoring setup in the Netherlands includes a "hospital-based telemonitoring center" with specialized e-nurses. Blood pressure data are automatically synchronized via the monitoring platform to a special health care provider dashboard, integrated into the electronic health record. The e-nurses in the telemonitoring center assess all the alarms generated by the monitoring platform based on the blood pressure data and discuss these alarms with clinicians if needed. A schematic overview of the HBPT processing steps is included in Multimedia Appendix 2. The clinicians supervising the e-nurses are internal medicine specialists, residents, or nurse practitioners who would also be involved in the SOC for patients with hypertension. They are also responsible for remotely adjusting blood pressure medication if needed.

## Model Input Parameters

### *Probabilities and Efficacy Input*

Multimedia Appendix 1 provides an overview of the baseline blood pressure distribution, annual event probabilities, and efficacy model inputs. Each systolic blood pressure state corresponded to a risk of a CV event, which was based on a large prospective real-world study [19]. The transition probability from a higher blood pressure state to a lower blood pressure state was based on a decrease of 5.1 mm Hg per year, which corresponded to the clinical effect of pharmacological therapy reported in the latest available meta-analysis [20] and applied to both groups. Patients in the HBPT group had an additional decrease of 12 mm Hg in systolic blood pressure in the first year due to HBPT. This additional effect was based on the latest available literature on the clinical effectiveness of HBPT [8]. A notable proportion, 19.7% of the patients had resistant hypertension resulting in the absence of blood pressure reduction [21]. The probability of dying from a CV event (CVD-related death) was based on the Blood Pressure Lowering Treatment Trialists' Collaboration study [18]. The probability of all-cause mortality was based on the age-based population mortality in the Netherlands [22] and was corrected for CVD-related deaths [22]. The probability of suffering a recurrent CV event was derived from a large study assessing the 10-year risk of recurrent vascular events [23].

### *Utilities*

The utilities of the model health states were derived from published literature (Multimedia Appendix 3 [27-30]). The baseline utility value was 0.96 for patients with hypertension [27], which declined to 0.79 following an MI [28], 0.64 following a cerebral infarction [28], and 0.59 following an intracranial hemorrhage [29]. The weighted average of these event utilities was 0.67 and was used in the model as the "postevent" utility for the year after the event occurred. The conservative assumption was made that after 1 year of the event, the utility would equal the baseline utility.

A recurrent intracerebral infarction or MI corresponded with a utility of 0.74 and 0.62, respectively [30]. For a recurrent intracranial hemorrhage, the utility value was considered equal to the utility value of a first intracranial hemorrhage, which was 0.59 [29]. The weighted average utility of a recurrence was 0.64 and was used for the year the recurrent event occurred and the subsequent years the patient was in the postrecurrent health state [27].

### *Costs and Discounting Rates*

Costs were divided into direct medical costs and nonmedical costs (Multimedia Appendix 4 [16,25,31-42]). For the HBPT group, direct medical costs consisted of a one-time out-of-pocket purchase of a blood pressure device [31], costs for remote monitoring [32], standard drug costs [33,43], additional drug costs [44], and in-person OPD consultations. The remote monitoring costs were based on an official Dutch tariff [32] for patients who are part of a remote monitoring program. A hospital can claim this tariff 3 times a year as a flat fee for a patient who is remotely monitored to cover costs for the license of telemonitoring software, salaries for the involved health care workers, and development costs. In the SOC group, direct medical costs only consisted of standard drug costs and costs for the in-person OPD consultations. Direct medical costs for a stroke (infarction and hemorrhage), MI, or CV-related death were based on data available from the Dutch National Health Care Institute and Ministry of Health, Welfare and Sport [34,35]. Total costs for each event were based on the overall reported expenses divided by the weighted incidences of both stroke and MI.

Nonmedical costs consisted of travel costs, parking costs, and costs related to productivity losses in both the SOC and HBPT groups. Productivity losses were based on work absence resulting from the in-person OPD consultations (1 hour for each visit) or due to an event (17.7 absent working days) and were based on data from the Dutch National Healthcare Institute [36] and the Trimbos Institute [37]. The costs of productivity losses were based on the average labor participation [38], average hourly wage [39], and average working week in the 65 - 75 years age group corresponding with the average age of 65.3 years used in the current analysis (based on the Blood Pressure Lowering Treatment Trialists' Collaboration study [18]. Friction costs following a death were calculated based on the friction costs method [40].

Discounting rates were 3% for the costs and 4% for the health outcomes based on the Dutch Economic Evaluation guidelines [15].

### Outcomes

The outcome measures used to compare the 2 interventions in this study were costs, QALYs, and incremental cost-effectiveness ratios (ICERs) presented as cost per QALY gained.

### Univariate Sensitivity Analysis and Scenario Analysis

To assess the impact of uncertainty on the ICER, an extensive sensitivity analysis was performed for the current early health economic analysis.

A univariate sensitivity analysis was performed to quantify the impact of parameter uncertainty on the ICER by varying all individual parameters one by one with ±20% of the mean. For utilities, the upper limit was restricted to a maximum of 1. In addition to the univariate sensitivity analysis, three scenario analyses were performed. (1) Since telemonitoring was expected to result in a reduction in the number of OPD consultations, the interdependency between these variables was assessed. The ICER was calculated for a range of telemonitoring costs and a range of frequencies of OPD consultations. (2) A scenario with a prolonged clinical effect of HBPT (2 and 3 years compared to 1 year in the base case) was modeled to assess the potential effect on the ICER. (3) To assess the impact of age on the ICER, the age at which remote patient monitoring is started was modeled over a range of 30 to 75 years.

### Assumptions

The following assumptions were made during model development. (1) A proportion of 19.7% of patients with hypertension was considered to have resistant hypertension [21]. About half of these patients have so-called “apparent resistant hypertension,” which is antihypertensive treatment failure due to drug nonadherence. We assumed that the HBPT intervention prevents nonadherence, resulting in only 9.85% of the patients having resistant hypertension in the HBPT group. (2) Patients receiving HBPT will have 50% fewer in-person consultations with their clinician or specialist nurse, as the remote patient monitoring partially replaces the need for in-hospital blood pressure measurements and identifies on-target patients who might not require a regular follow-up consultation. It was assumed that the patients in the HBPT group would on average have 1.5 in-person OPD consultations annually. (3) Since most of the HBPT trials have follow-up durations of up to 1 year, we assumed that HBPT would only cause an additional blood pressure lowering effect (in addition to the effect of drug therapy) in the first year (cycle 1). (4) HBPT prevents patients from suffering from overtreatment (blood pressure <120 mm Hg), which also results in an increased risk for CV events and death. Therefore, in the SOC group, patients could transition to the <120 mm Hg health state for a maximum of 1 year after which they returned to the 120 - 129 mm Hg health state. In the HBPT group, it was assumed that patients could not transition to the <120 mm Hg health state. (5) The second year after a CV event, patients will return back to the baseline utility.

### Ethical Considerations

No ethics approval was applied for this study as this study was not conducted on newly generated real-world data from human participants. Data for the probabilities, costs, and utilities were derived from the available literature or from publicly available government sources.

## Results

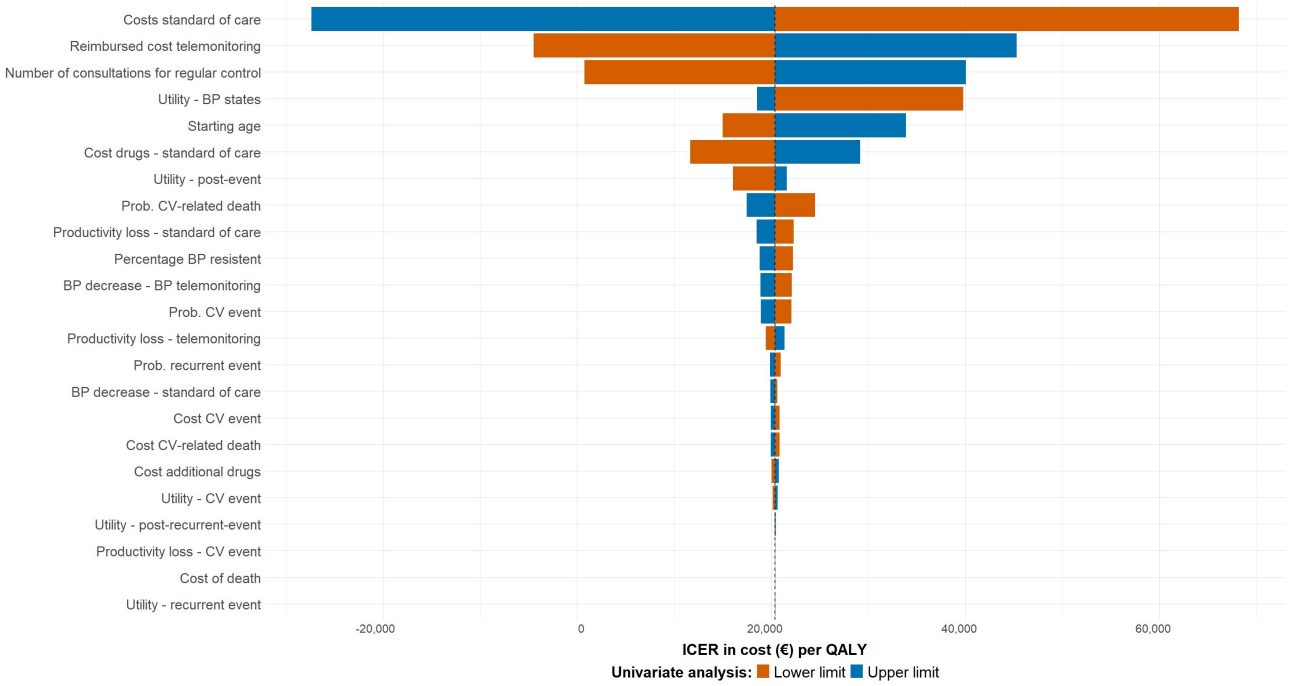
### Base Case

In the base case, the cost for the HBPT group was €20,463,881 and €19,196,847 in the SOC group, resulting in incremental costs for HBPT of €1,267,034 compared to SOC. Additionally, HBPT resulted in 13,401.19 QALYs compared to 13,339.04 QALYs in the SOC group, resulting in an incremental effect of 62.15 QALYs in favor of HBPT. The resulting ICER for the base-case analysis was €20,386 per QALY. Based on the WTP threshold of €20,000 per QALY [14], telemonitoring is not considered cost-effective following the assumptions of the base-case analysis. The additional costs of telemonitoring outweigh the QALYs gained because of prevented first and recurrent CV events.

### Univariate Sensitivity Analysis

The results of the univariate sensitivity analysis indicate the impact of parameter uncertainty on the ICER. Based on the results in Figure 2, the uncertainty in the additional costs of SOC resulting from in-person OPD consultations has the highest impact on the ICER. In case the upper limit was selected for the SOC costs (€1166.33), HBPT became cost-saving compared to SOC. With the lower limit of the reimbursed costs for telemonitoring (€403.20 instead of €504.00 in the base case), HBPT also became cost-saving compared to SOC. In case the number of consultations was reduced to 1.2 per year in the HBPT group, the ICER dropped to €742 per QALY and HBPT was considered cost-effective. In all other cases, the analysis of parameter uncertainty resulted in an ICER between €11,641 per QALY and €39,758 per QALY.

**Figure 2.** Tornado diagram of the results of the univariate sensitivity analysis (€1=US $1.09). BP: blood pressure; CV: cardiovascular; ICER: incremental cost-effectiveness ratio; Prob: probability; QALY: quality-adjusted life year.

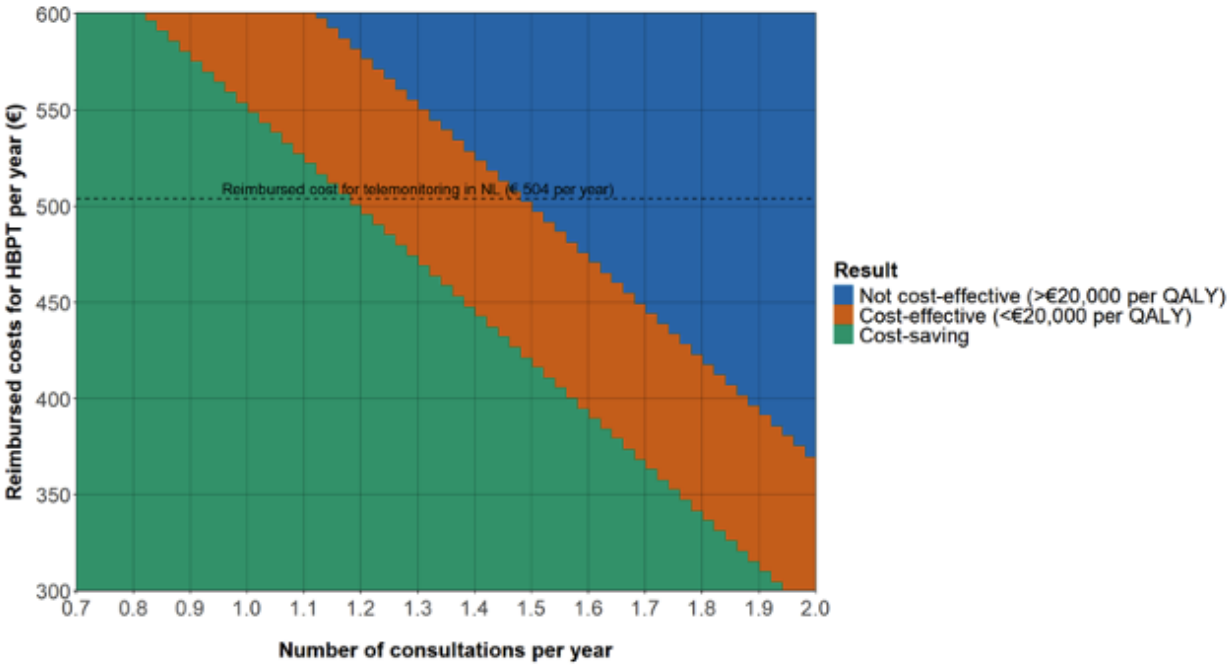



## Scenario Analysis

### *Scenario 1: Variable Telemonitoring Costs and Frequency of OPD Consultations*

One important assumption of the current model is related to the number of in-person consultations for patients in the HBPT group. It was assumed that in the HBPT group, the number of annual consultations dropped from 3 to 1.5 per year. A further decrease in the number of consultations could result in a further reduction of the ICER. Based on the results of scenario 1, HBPT will become cost-effective (<€20,000 per QALY) with the current reimbursement of €504 per year at 1.48 in-person OPD consultations per year and will become cost-saving at 1.18 in-person OPD consultations per year (Figure 3).

**Figure 3.** ICER results (not cost-effective, cost-effective, and cost-saving) of scenario analysis calculated over a range of costs for HBPT per year and a range of in-person consultations per year (€1=US $1.09). HBPT: home blood pressure telemonitoring; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year.



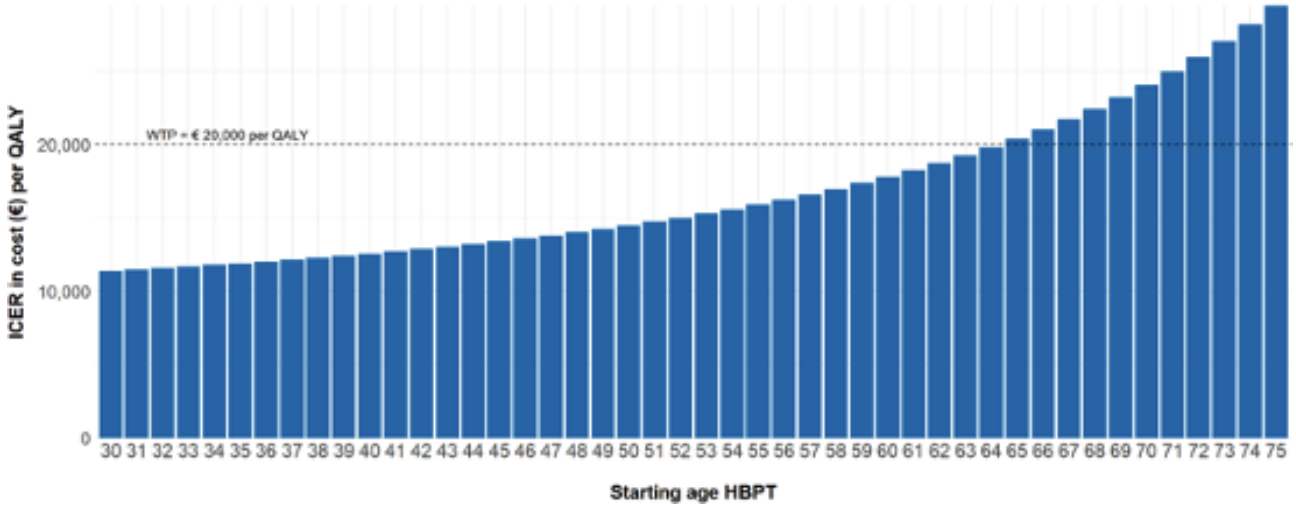
### *Scenario 2: Prolonged Clinical Effect of HBPT*

The duration of the effect of HBPT comes with great uncertainty and was assumed to last for only 1 year in the base-case analysis, which could be considered conservative. Scenario 2 indicates that the ICER will reach €11,154 per QALY in case the effect of HBPT lasts for 2 years (a total blood pressure reduction of 24 mm Hg after 2 years) and further declines to €9204 per QALY in case HBPT reduces the blood pressure with 12 mm Hg for 3 years (a total blood pressure reduction of 36 mm Hg after 3 years).

### *Scenario 3: Variable Starting Age HBPT*

The results of scenario 3, in which the model was run over an age range of 30 to 75 years (Figure 4), indicate that the younger the patient's age at the start of HBPT, the lower the ICER. If HBPT is started at the age of 64 years or below, HBPT could be considered a cost-effective intervention.

**Figure 4.** ICER per age at which HBPT is started (€1=US $1.09). HBPT: home blood pressure telemonitoring; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year; WTP: willingness to pay.




## Discussion

### Principal Findings

#### *Overview*

To the best of our knowledge, this is the first study to provide an early cost-effectiveness analysis of HBPT in patients with hypertension without previous CV events. Based on the current early cost-effectiveness model, which reflects a societal perspective and includes both short- and long-term costs and benefits, HBPT showed the potential to be cost-effective following realistic reductions in SOC for patients with hypertension. Specifically, a reduction in the number of OPD consultations in the HBPT group will make HBPT cost-effective or even cost-saving. These findings underscore the potential of HBPT and with that the importance of genuine digital transformation in health care, advocating for the substitution of traditional OPD care with digital and remote care, rather than providing digital care as an add-on to standard OPD care. Additionally, we found that early (ie, younger age) and sustained telemonitoring further improves the cost-effectiveness of HBPT.

#### *Outcome-Based Health Care*

Policy makers should make use of the fact that HBPT can be cost-effective following a reduction in the number of OPD consultations. The current reimbursement structure provided by the Dutch Health Care Authority [32] does not include any requirements in terms of reducing standard care and therefore seems unsuitable in its current form. Outcome-based health care contracts [45], characterized by performance fees linked to predefined shared objectives between hospitals and insurance companies, are ideally suited for HBPT to pursue genuine digital transformation efforts. By setting a shared objective in terms of physical care replacement with HBPT, a sustainable system could be established that allows for efficient (with fewer resources) and cost-effective care delivery along the lines of value chain optimization and significant displacement effects.

#### *The Current Business Model for HBPT*

Hypothetically, HBPT should be able to realize short-term benefits through greater efficiency with regard to care organization and long-term benefits, resulting from a greater level of blood pressure control, which translates into a reduction in CV events and CV-related deaths. We found that short-term benefits realized through a substantial reduction in OPD consultations had a major impact on the ICER, but the impact of long-term benefits appeared to be limited. The 1-year effect of HBPT on the blood pressure of patients resulted in minor between-group differences in CV events or CVD-related deaths. Extending the effect of HBPT to 2 or 3 years substantially reduced the ICER, but clinical evidence supporting this assumption is lacking, as follow-up durations in clinical trials are usually no longer than 12 months [8]. Therefore, short-term benefits resulting from more efficient care delivery are expected to become the major driver for a sustainable business model for HBPT. The further upward potential of remote monitoring comes with a multimorbidity perspective (eg, hypertension and diabetes). Many patients have a variety of comorbidities and multiple consultations with different specialists. If one remote monitoring program reduces the number of consultations across multiple medical specialties, remote monitoring is more likely to result in cost savings.

### *Alternative Models Available in the Literature*

The only available comparable study [46] evaluates the cost-effectiveness of HBPT in a poststroke population using a Markov cohort simulation. The HBPT intervention was cost-saving in the base case and cost-effective in the scenario analyses with an ICER of US $1200-4700 per QALY. In contrast to our study, the benefit in terms of blood pressure reduction due to HBPT was modeled as a continuous effect (year after year). Even though our scenario analyses with 2 and 3 years of clinical benefits of HBPT resulted in HBPT being cost-effective, the results of the previously described study [46] should be considered as optimistic as evidence on a sustained (year after year) effect is lacking. Other studies that reported a positive effect of remote monitoring include heart failure monitoring [47-49] or monitoring of patients with COVID-19 [50]. These studies [46,47,50] highlight the importance of reducing short-term care consumption to come to a favorable ICER. This advocates for the substitution of traditional care with digital and remote care, rather than providing digital care as an add-on to standard care.

## Limitations

The current early health economic analysis comes with limitations that should be considered when interpreting the results of this study. First, we did not consider any effects of HBPT on diastolic blood pressure, which could have resulted in potential CVD risk reduction. However, since most of the hypertensive population has systolic or both systolic and diastolic hypertension, the impact of this simplification is expected to be limited.

Second, baseline blood pressure distribution and CVD-related death were derived from one study, which could impact the generalizability of the results [19]. However, given the large number of patients included in the study (n=96,268) and the follow-up period of 10 years, the study was considered highly valuable for the current early health economic analysis. Gathering country-specific data on the baseline blood pressure distribution and CVD-related death will become important when the current model is applied to inform reimbursement decisions.

Third, as many patients with hypertension often have other relevant comorbidities, reducing the number of in-person visits during the HBPT program might negatively impact the provided care and cost-effectiveness for other relevant diseases, which would normally be addressed during the same consultation. It appears to be more likely, however, that future remote patient monitoring programs will encompass multiple conditions (eg, hypertension and diabetes) and thereby overcome this potential disadvantage.

Fourth, the model does not allow patients to move to higher systolic blood pressure states, which could result in an overestimation of long-term blood pressure regulation. This limitation was partially overcome by classifying part of the population as apparently resistant, implying their blood pressure did not decrease. Furthermore, any potential overestimation would affect both the SOC and HBPT groups, thus largely neutralizing the impact on comparative results.

Future research should focus on reducing uncertainty on key input parameters, which include the duration of the effect and the number of OPD consultations per year needed in addition to HBPT. Additionally, future research should focus on the effect of scale in terms of the number of patients included in the HBPT program as an additional prerequisite for sustainable implementation, as the one-time investment costs are substantial when starting with HBPT. Moreover, this model should be validated with real-world data, specifically from a Dutch randomized trial. Finally, future research should consider the cost-effectiveness across different care settings, as a significant portion of patients with hypertension are treated by general practitioners.

## Conclusion

Based on the current early health economic analysis, we found HBPT to be cost-effective, provided it will result in a genuine digital transformation in health care and thereby substantially reduce the number of standard OPD consultations.

## Data Availability

The R model is available from the authors upon reasonable request.

## Conflicts of Interest

DD is the CEO of Luscii, the home blood pressure telemonitoring (HBPT) platform, which is used in the HBPT program that was analyzed in this study. There was no financial support provided by Luscii to facilitate this study.

Multimedia Appendix 1
Baseline blood pressure distribution, annual event probabilities, and efficacy model inputs.
[DOCX File, 30 KB - cardio_v9i1e64386_app1.docx ]

Multimedia Appendix 2
Telemonitoring organization and alert processing.
[DOCX File, 143 KB - cardio_v9i1e64386_app2.docx ]

Multimedia Appendix 3
Health utility values.
[DOCX File, 25 KB - cardio_v9i1e64386_app3.docx ]

Multimedia Appendix 4
Costs and discounting rates.
[DOCX File, 30 KB - cardio_v9i1e64386_app4.docx ]

Checklist 1
CHEERS (Consolidated Health Economic Evaluation Reporting Standards) 2022 checklist.
[DOCX File, 27 KB - cardio_v9i1e64386_app5.docx ]

## References


1. Fuchs FD, Whelton PK. High blood pressure and cardiovascular disease. Hypertension 2020 Feb;75(2):285-292. [doi: 10.1161/HYPERTENSIONAHA.119.14240] [Medline: 31865786]
2. Schelleman H, Klungel OH, Kromhout D, de Boer A, Stricker BHC, Verschuren WMM. Prevalence and determinants of undertreatment of hypertension in the Netherlands. J Hum Hypertens 2004 May;18(5):317-324. [doi: 10.1038/sj.jhh.1001672] [Medline: 15103311]
3. Duan Y, Xie Z, Dong F, et al. Effectiveness of home blood pressure telemonitoring: a systematic review and meta-analysis of randomised controlled studies. J Hum Hypertens 2017 Jul;31(7):427-437. [doi: 10.1038/jhh.2016.99] [Medline: 28332506]
4. Versmissen J, van Steenkiste J, Koch BCP, Peeters LEJ. “Under pressure”: the role of therapeutic drug monitoring in the treatment of hypertension. Br J Clin Pharmacol 2024 Aug;90(8):1884-1891. [doi: 10.1111/bcp.16125] [Medline: 38845455]
5. Parati G, Dolan E, McManus RJ, Omboni S. Home blood pressure telemonitoring in the 21st century. J Clin Hypertens (Greenwich) 2018 Jul;20(7):1128-1132. [doi: 10.1111/jch.13305] [Medline: 30003701]
6. Luscii. URL: https://luscii.com/en/home [accessed 2024-7-12]
7. Patient Journey App. URL: https://patientjourneyapp.com [accessed 2024-7-12]
8. Kaihara T, Intan-Goey V, Scherrenberg M, et al. Automatic transmission of home blood pressure data can be effective in managing hypertension: a systematic review and meta-analysis. Eur Heart J Digit Health 2022 Dec;3(4):638-653. [doi: 10.1093/ehjdh/ztac049] [Medline: 36710899]
9. Gijsbers H, Feenstra TM, Eminovic N, et al. Enablers and barriers in upscaling telemonitoring across geographic boundaries: a scoping review. BMJ Open 2022 Apr 20;12(4):e057494. [doi: 10.1136/bmjopen-2021-057494] [Medline: 35443957]
10. Dehmer SP, Maciosek MV, Trower NK, et al. Economic evaluation of the home blood pressure telemonitoring and pharmacist case management to control hypertension (Hyperlink) trial. J Am Coll Clin Pharm 2018 Oct;1(1):21-30. [doi: 10.1002/jac5.1001] [Medline: 30320302]
11. Stoddart A, Hanley J, Wild S, et al. Telemonitoring-based service redesign for the management of uncontrolled hypertension (HITS): cost and cost-effectiveness analysis of a randomised controlled trial. BMJ Open 2013 May 28;3(5):e002681. [doi: 10.1136/bmjopen-2013-002681] [Medline: 23793650]
12. Omboni S, Gazzola T, Carabelli G, Parati G. Clinical usefulness and cost effectiveness of home blood pressure telemonitoring: meta-analysis of randomized controlled studies. J Hypertens 2013 Mar;31(3):455-467. [doi: 10.1097/HJH.0b013e32835ca8dd] [Medline: 23299557]
13. Yee M, Tappenden P, Wailoo A. NICE DSU report: economic evaluation in NICE early value assessments. : The University of Sheffield; 2023 URL: https://www.sheffield.ac.uk/nice-dsu/methods-development/eva [accessed 2025-04-15]
14. Boersma C, Broere A, Postma MJ. Quantification of the potential impact of cost-effectiveness thresholds on Dutch drug expenditures using retrospective analysis. Value Health 2010;13(6):853-856. [doi: 10.1111/j.1524-4733.2010.00736.x] [Medline: 21054659]
15. Guideline for conducting health economic evaluations [Report in Dutch]. : Dutch Healthcare Institute; 2024 URL: https://www.zorginstituutnederland.nl/publicaties/publicatie/2024/01/16/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg [accessed 2024-07-12]
16. Consumer price indexation [Web page in Dutch]. Central Statistics Office. URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70936NED/table?fromstatweb [accessed 2024-07-18]
17. R: a language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.R-project.org [accessed 2025-04-15]
18. Blood Pressure Lowering Treatment Trialists’ Collaboration. Pharmacological blood pressure lowering for primary and secondary prevention of cardiovascular disease across different levels of blood pressure: an individual participant-level data meta-analysis. Lancet 2021 May 1;397(10285):1625-1636. [doi: 10.1016/S0140-6736(21)00590-0] [Medline: 33933205]
19. Peng X, Jin C, Song Q, Wu S, Cai J. Stage 1 hypertension and the 10-year and lifetime risk of cardiovascular disease: a prospective real-world study. J Am Heart Assoc 2023 Apr 4;12(7):e028762. [doi: 10.1161/JAHA.122.028762] [Medline: 36975094]

20. Canoy D, Copland E, Nazarzadeh M, et al. Antihypertensive drug effects on long-term blood pressure: an individual-level data meta-analysis of randomised clinical trials. Heart 2022 Jul 27;108(16):1281-1289. [doi: 10.1136/heartjnl-2021-320171] [Medline: 35058294]
21. Carey RM, Sakhuja S, Calhoun DA, Whelton PK, Muntner P. Prevalence of apparent treatment-resistant hypertension in the United States. Hypertension 2019 Feb;73(2):424-431. [doi: 10.1161/HYPERTENSIONAHA.118.12191] [Medline: 30580690]
22. Age-based mortality [Web page in Dutch]. Central Statistics Office. URL: https://www.cbs.nl/nl-nl/cijfers/detail/37360ned [accessed 2024-07-12]
23. Kaasenbrood L, Boekholdt SM, van der Graaf Y, et al. Distribution of estimated 10-year risk of recurrent vascular events and residual risk in a secondary prevention population. Circulation 2016 Nov 8;134(19):1419-1429. [doi: 10.1161/CIRCULATIONAHA.116.021314] [Medline: 27682883]
24. Mancia G, Kreutz R, Brunström M, et al. 2023 ESH Guidelines for the management of arterial hypertension The Task Force for the management of arterial hypertension of the European Society of Hypertension: endorsed by the International Society of Hypertension (ISH) and the European Renal Association (ERA). J Hypertens 2023 Dec 1;41(12):1874-2071. [doi: 10.1097/HJH.0000000000003480] [Medline: 37345492]
25. Dbc-pakket 2024 rz24a [Web page in Dutch]. Dutch Healthcare Authority. 2024 Jul 12. URL: https://puc.overheid.nl/nza/doc/PUC_740006_22/1/ [accessed 2025-04-16]
26. Telemonitoring and e-coaching in hypertension. ClinicalTrials.gov. URL: https://clinicaltrials.gov/study/NCT05660226 [accessed 2024-07-12]
27. Xie X, He T, Kang J, Siscovick DS, Li Y, Pagán JA. Cost-effectiveness analysis of intensive hypertension control in China. Prev Med 2018 Jun;111:110-114. [doi: 10.1016/j.ypmed.2018.02.033] [Medline: 29525577]
28. Betts MB, Rane P, Bergrath E, et al. Utility value estimates in cardiovascular disease and the effect of changing elicitation methods: a systematic literature review. Health Qual Life Outcomes 2020 Jul 27;18(1):251. [doi: 10.1186/s12955-020-01407-y] [Medline: 32718355]
29. Sallinen H, Sairanen T, Strbian D. Quality of life and depression 3 months after intracerebral hemorrhage. Brain Behav 2019 May;9(5):e01270. [doi: 10.1002/brb3.1270] [Medline: 30907075]
30. Zomer E, Si S, Hird TR, et al. Cost-effectiveness of low-dose rivaroxaban and aspirin versus aspirin alone in people with peripheral or carotid artery disease: an Australian healthcare perspective. Eur J Prev Cardiol 2019 May;26(8):858-868. [doi: 10.1177/2047487318817910] [Medline: 30526023]
31. IHealth track blood pressure machine [Web page in Dutch]. Bol. URL: https://www.bol.com/nl/nl/p/[illegible]9200000061056044/?bltgh=tXYw2la61G4fI9ZABXRFBQ.4_27.28.ProductImage [accessed 2024-07-12]
32. Digital health reimbursement 2024 [Web page in Dutch]. Dutch Healthcare Authority. URL: https://puc.overheid.nl/nza/doc/PUC_655318_22/1 [accessed 2024-07-12]
33. Pharmaceutical care integral funding [Report in Dutch]. : Netherlands Institute for Health Services Research; 2011 URL: https://www.nivel.nl/sites/default/files/bestanden/Rapport-Farmaceutische-zorg-integrale-bekostiging.pdf [accessed 2024-07-12]
34. van Eeden M, van Heugten C, van Mastrigt GAPG, van Mierlo M, Visser-Meily JMA, Evers SMAA. The burden of stroke in the Netherlands: estimating quality of life and costs for 1 year poststroke. BMJ Open 2015 Nov 27;5(11):e008220. [doi: 10.1136/bmjopen-2015-008220] [Medline: 26614618]
35. Soekhlal RR, Burgers LT, Redekop WK, Tan SS. Treatment costs of acute myocardial infarction in the Netherlands. Neth Heart J 2013 May;21(5):230-235. [doi: 10.1007/s12471-013-0386-y] [Medline: 23456884]
36. Cost manual for economic evaluations in healthcare: methodology and reference prices [Web page in Dutch]. Dutch Healthcare Institute. URL: https://www.zorginstituutnederland.nl/over-ons/werkwijzen-en-procedures/adviseren-over-en-verduidelijken-van-het-basispakket-aan-zorg/beoordeling-van-geneesmiddelen/richtlijn-voor-economische-evaluatie [accessed 2024-07-12]
37. Absence due to physical conditions in a working population [Web page in Dutch]. Trimbos Institute. URL: https://www.trimbos.nl/docs/af1092-verzuim-door-psychische-en-somatische-aandoeningen-bij-werkenden.pdf [accessed 2024-07-12]
38. Labor participation in the Netherlands [Web page in Dutch]. Central Statistics Office. URL: https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/werkenden#:~:text=De%20werkzame%20beroepsbevolking%20bestaat%20uit,in%20het%20buitenland%20werkzaam%20zijn [accessed 2024-07-12]
39. Average hourly wage in the Netherlands [Web page in Dutch]. Central Statistics Office. URL: https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/ontwikkeling-cao-lonen/uurloon#:~:text=Als%20we%20kijken%20naar%20gemiddelden,van%2028%20euro%20of%20hoger [accessed 2024-07-12]
40. Quality of health measurements [Web page in Dutch]. Dutch Healthcare Institute. URL: https://www.zorginstituutnederland.nl/binaries/zinl/documenten/publicatie/2024/01/16/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg/Verdiepingsmodule+QALY+en+kwaliteit-van-leven-metingen+%28versie+2024%29.pdf [accessed 2024-07-12]

41. DHI amlodipine drug costs [Web page in Dutch]. Dutch Healthcare Institute. URL: https://www.medicijnkosten.nl/zoeken?trefwoord=amlodipine [accessed 2024-07-18]
42. Average working week in the Netherlands [Web page in Dutch]. Central Statistics Office. URL: https://www.cbs.nl/nl-nl/nieuws/2020/08/meer-dan-de-helft-werkt-voltijds [accessed 2024-07-12]
43. Guideline for the treatment of patients with severe hypertension [Web page in Dutch]. KvK Federation of Medical Specialists. URL: https://richtlijnendatabase.nl/richtlijn/hypertensief_spoedgeval/behandeling_van_ernstige_hypertensie_zonder_acute_orgaanschade.html [accessed 2024-07-12]
44. McManus RJ, Mant J, Haque MS, et al. Effect of self-monitoring and medication self-titration on systolic blood pressure in hypertensive patients at high risk of cardiovascular disease: the TASMIN-SR randomized clinical trial. JAMA 2014 Aug 27;312(8):799-808. [doi: 10.1001/jama.2014.10057] [Medline: 25157723]
45. Bohm N, Bermingham S, Grimsey Jones F, et al. The challenges of outcomes-based contract implementation for medicines in Europe. Pharmacoeconomics 2022 Jan;40(1):13-29. [doi: 10.1007/s40273-021-01070-1] [Medline: 34480324]
46. Padwal RS, So H, Wood PW, et al. Cost-effectiveness of home blood pressure telemonitoring and case management in the secondary prevention of cerebrovascular disease in Canada. J Clin Hypertens (Greenwich) 2019 Feb;21(2):159-168. [doi: 10.1111/jch.13459] [Medline: 30570200]
47. Albuquerque de Almeida F, Corro Ramos I, Al M, Rutten-van Mölken M. Home telemonitoring and a diagnostic algorithm in the management of heart failure in the Netherlands: cost-effectiveness analysis. JMIR Cardio 2022 Aug 4;6(2):e31302. [doi: 10.2196/31302] [Medline: 35925670]
48. Kokkonen J, Mustonen P, Heikkilä E, et al. Effectiveness of telemonitoring in reducing hospitalization and associated costs for patients with heart failure in Finland: nonrandomized pre-post telemonitoring study. JMIR Mhealth Uhealth 2024 Feb 7;12:e51841. [doi: 10.2196/51841] [Medline: 38324366]
49. Jiang X, Yao J, You JHS. Cost-effectiveness of a telemonitoring program for patients with heart failure during the COVID-19 pandemic in Hong Kong: model development and data analysis. J Med Internet Res 2021 Mar 3;23(3):e26516. [doi: 10.2196/26516] [Medline: 33656440]
50. van Herwerden MC, van Steenkiste J, El Moussaoui R, den Hollander JG, Helfrich G, J A M Verberk I. Home telemonitoring and oxygen therapy in COVID-19 patients: safety, patient satisfaction, and cost-effectiveness [Article in Dutch]. Ned Tijdschr Geneeskd 2021 Mar 11;165:D5740. [Medline: 33720552]


## Abbreviations

**CHEERS:** Consolidated Health Economic Evaluation Reporting Standards
**CV:** cardiovascular
**CVD:** cardiovascular disease
**HBPT:** home blood pressure telemonitoring
**ICER:** incremental cost-effectiveness ratio
**MI:** myocardial infarction
**OPD:** outpatient department
**QALY:** quality-adjusted life year
**SOC:** standard of care
**WTP:** willingness to pay


*Edited by A Coristine; submitted 18.07.24; peer-reviewed by A Stanimirovic, E Khoong, J Edwards, MY Wu; revised version received 04.02.25; accepted 11.02.25; published 08.05.25.*

*Please cite as:*
*van Steenkiste J, van Dorst P, Dohmen D, Boersma C*
*Prerequisites for Cost-Effective Home Blood Pressure Telemonitoring: Early Health Economic Analysis*
*JMIR Cardio 2025;9:e64386*
*URL: https://cardio.jmir.org/2025/1/e64386*
*doi:10.2196/64386*

# Efficiency Improvement of the Clinical Pathway in Cardiac Monitor Insertion and Follow-Up: Retrospective Analysis

Ville Vanhala[1], MD; Outi Surakka[2], MHA; Vilma Multisilta[1], BSN; Mette Lundsby Johansen[3], MSci; Jonas Villinger[4], MD; Emmanuelle Nicolle[4], MD; Johanna Heikkilä[2], PhD; Pentti Korhonen[1], MD

[1]Tampere Heart Hospital, Elämänaukio 1, Tampere, Finland

[2]Jamk University of Applied Sciences, Jyväskylä, Finland

[3]Medtronic Denmark, Copenhagen, Denmark

[4]Medtronic International Trading Sarl, Tolochenaz, Switzerland

**Corresponding Author:**
Ville Vanhala, MD
Tampere Heart Hospital, Elämänaukio 1, Tampere, Finland

## Abstract

**Background:** The insertable cardiac monitor (ICM) clinical pathway in Tampere Heart Hospital, Finland, did not correspond to the diagnostic needs of the population. There has been growing evidence of delegating the insertion from cardiologists to specially trained nurses and outsourcing the remote follow-up. However, it is unclear if the change in the clinical pathway is safe and improves efficiency.

**Objective:** We aim to describe and assess the efficiency of the change in the ICM clinical pathway.

**Methods:** Pathway improvements included initiating nurse-performed insertions, relocating the procedure from the catheterization laboratory to a procedure room, and outsourcing part of the remote follow-up to manage ICM workload. Data were collected from electronic health records of all patients who received an ICM in the Tampere Heart Hospital in 2018 and 2020. Follow-up time was 36 months after insertion.

**Results:** The number of inserted ICMs doubled from 74 in 2018 to 159 in 2020. In 2018, cardiologists completed all insertions, while in 2020, a total of 70.4% (n=112) were completed by nurses. The waiting time from referral to procedure was significantly shorter in 2020 (mean 36, SD 27.7 days) compared with 2018 (mean 49, SD 37.3 days; *P*=.02). The scheduled ICM procedure time decreased from 60 minutes in 2018 to 45 minutes in 2020. Insertions performed in the catheterization laboratory decreased significantly (n=14, 18.9% in 2018 and n=3, 1.9% in 2020; *P*=<.001). Patients receiving an ICM after syncope increased from 71 to 94 patients. Stroke and transient ischemic attack as an indication increased substantially from 2018 to 2020 (2 and 62 patients, respectively). In 2018, nurses analyzed all remote transmissions. In 2020, the external monitoring service escalated only 11.2% (204/1817) of the transmissions to the clinic for revision. This saved 296 hours of nursing time in 2020. Having nurses insert ICMs in 2020 saved 48 hours of physicians' time and the shorter scheduling for the procedure saved an additional 40 hours of nursing time compared with the process in 2018. Additionally, the catheterization laboratory was released for other procedures (27 h/y). The complication rate did not change significantly (n=2, 2.7% in 2018 and n=5, 3.1% in 2020; *P*=.85). The 36-month diagnostic yield for syncope remained high in 2018 and 2020 (n=32, 45.1% and n=36, 38.3%; *P*=.38). The diagnostic yield for patients who had stroke with a procedure in 2020 was 43.5% (n=27).

**Conclusions:** The efficiency of the clinical pathway for patients eligible for an ICM insertation can be increased significantly by shifting to nurse-led insertions in procedure rooms and to the use of an external monitoring and triaging service.

***(JMIR Cardio 2025;9:e67774)*** doi:10.2196/67774

**KEYWORDS**

insertable cardiac monitor; clinical pathway; nurse-led service; task shifting; efficiency improvement; remote monitoring

## Introduction

### Background

Insertable cardiac monitors (ICMs) are indicated for long-term monitoring of heart rhythms, primarily for the indications of unexplained syncope and cryptogenic stroke (CS) or transient ischemic attack (TIA) [1-4]. For patients monitored with an ICM, a remote monitoring system transfers ICM data daily to the hospital staff for analysis. The 2023 European Heart Rhythm Association–Heart Rhythm Society expert consensus on remote monitoring recommends remote monitoring as standard of care for ICMs [5]. However, remote monitoring can create a significant data burden [6], which can be challenging in the current context of clinical staff shortage and disparities between

different populations for access to services [7]. Recent studies have indicated that the in-office time to follow-up an ICM patient took approximately 39.9 minutes of staff time, while remote follow-up required only 11.3 minutes [8]. In addition, in studies regarding nurse-led ICM service, it has been confirmed that in an outpatient setting, ICM service by specially trained nurses can lead to significant savings without compromising the safety of the procedure [6].

Workforce challenges are well-known across countries. Therefore, the 2023 European Heart Rhythm Association–Heart Rhythm Society consensus statement recommends the effective management of remote monitoring clinics to focus on adequate staffing with clear roles and responsibilities, on-going staff education, and efficient high-priority alert systems [5]. Nurse-led services play a particularly important role for efficient ICM services, as international case studies show that nurses can conduct both ICM insertions and remote follow-up effectively and safely [9].

Additionally, the use of third-party resources can be an opportunity to efficiently manage remote monitoring of ICM patients and a solution for dealing with increased device clinic volume [8,10]. ICMs are prone to produce a heavy workload for the remote monitoring clinic (25% of all transmissions, 10 times more frequent than for a pacemaker) [11].

In Finland, health services are challenged due to the shortage of trained health care professionals and resources. For example, Finland has fewer cardiologists than the average for the member countries of the European Society of Cardiology (ESC; Finland 50.5 per million people vs ESC countries 85.1 per million people) [7]. Finland also faces a growing need for nurses in Finland [12]. The Finnish government has launched the “Good Work Program” to ensure the sufficiency and availability of personnel in health care, social welfare, and rescue services. The program aims to increase the attractiveness of working within the social and health care sector by developing the structures and clarifying the tasks between the personnel [13].

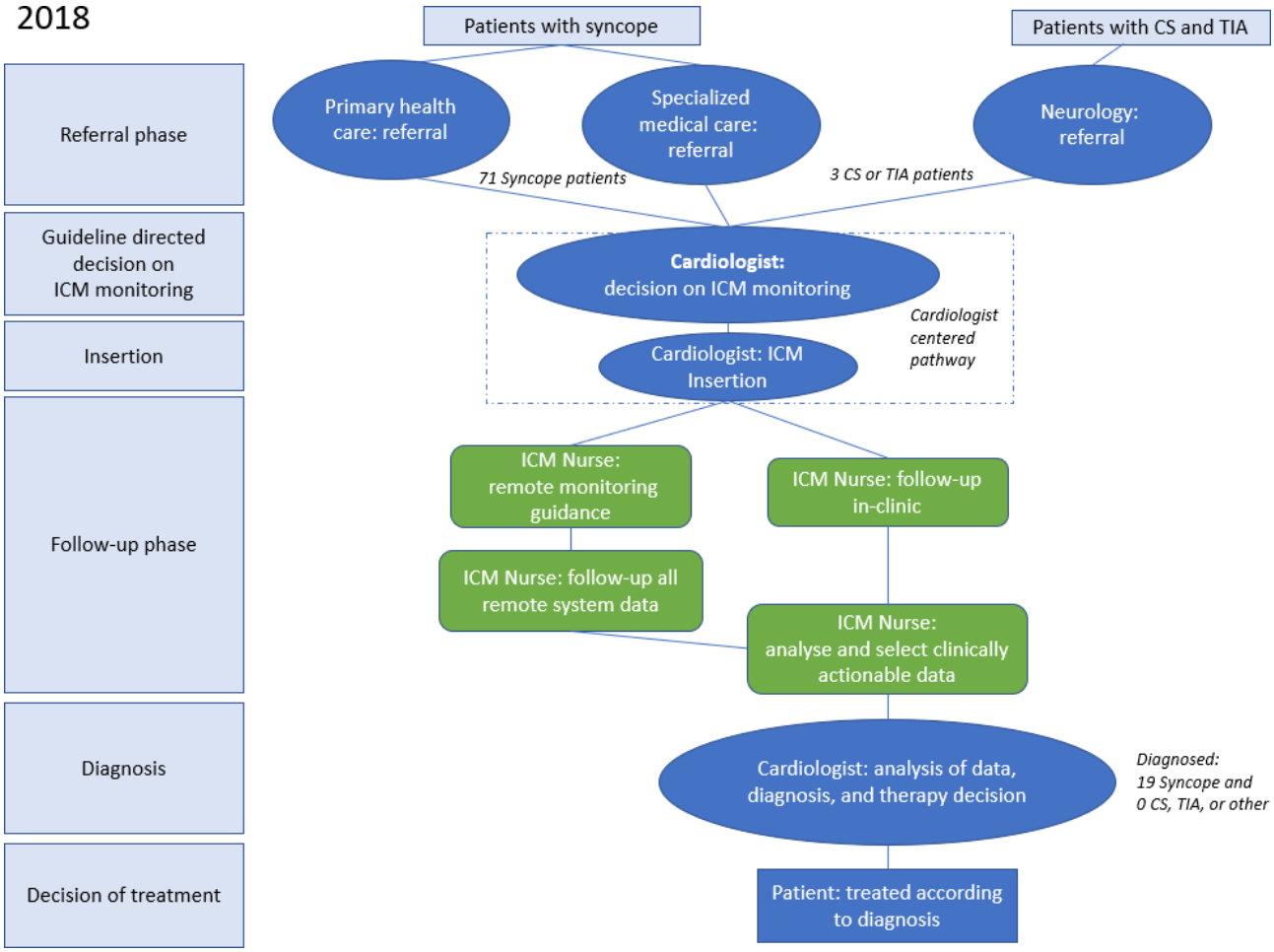
At the Finnish Tampere Heart Hospital, both insufficient staff resources and a growing number of patients in need of ICM monitoring led to the restructuring of the clinical patient pathway. The changes centered around training nurses to perform ICM insertions, the inclusion of the neurology department in patient pathways, moving the remaining ICM procedures out of the catheter laboratory, and the use of third-party triaging services.

However, the impact of these changes from the perspective of efficient resource management and quality of care is unknown. Thus, we conducted an analysis of the changes in clinical pathways at the Tampere Heart Hospital, assessing the impact on patient pathway efficiencies and the quality of care.

## Analyzing the ICM Pathway in 2018

In 2018, the Tampere Heart Hospital analyzed the prevailing ICM clinical pathway, and the way tasks were divided between professionals in each phase. The 2018 patient pathway was characterized by cardiology-centric decision-making for ICM insertions. Only a few patients who had CS were referred to the cardiology department even though the neurologist could make a referral to atrial fibrillation (AF) monitoring therapy for secondary prevention of CS and TIA. At the time, the ESC guidelines for AF management from 2016 were valid [3]. Unexplained patients who had syncope were referred by a general practitioner or the emergency department doctor to a cardiology clinic, where a cardiologist assessed whether these patients required an ICM based on the ESC guidelines from 2018 [1]. If an ICM was recommended for CS, TIA, or unexplained syncope, the patient was placed on a waiting list for the procedure and later invited to an outpatient clinic for device insertion by a cardiologist in a catheterization laboratory (Figure 1). The laboratory time was a highly demanded resource for performing more advanced interventional cardiological procedures.

**Figure 1.** Patient pathways in 2018. CS: cryptogenic stroke; ICM: insertable cardiac monitor; TIA: transient ischemic attack.

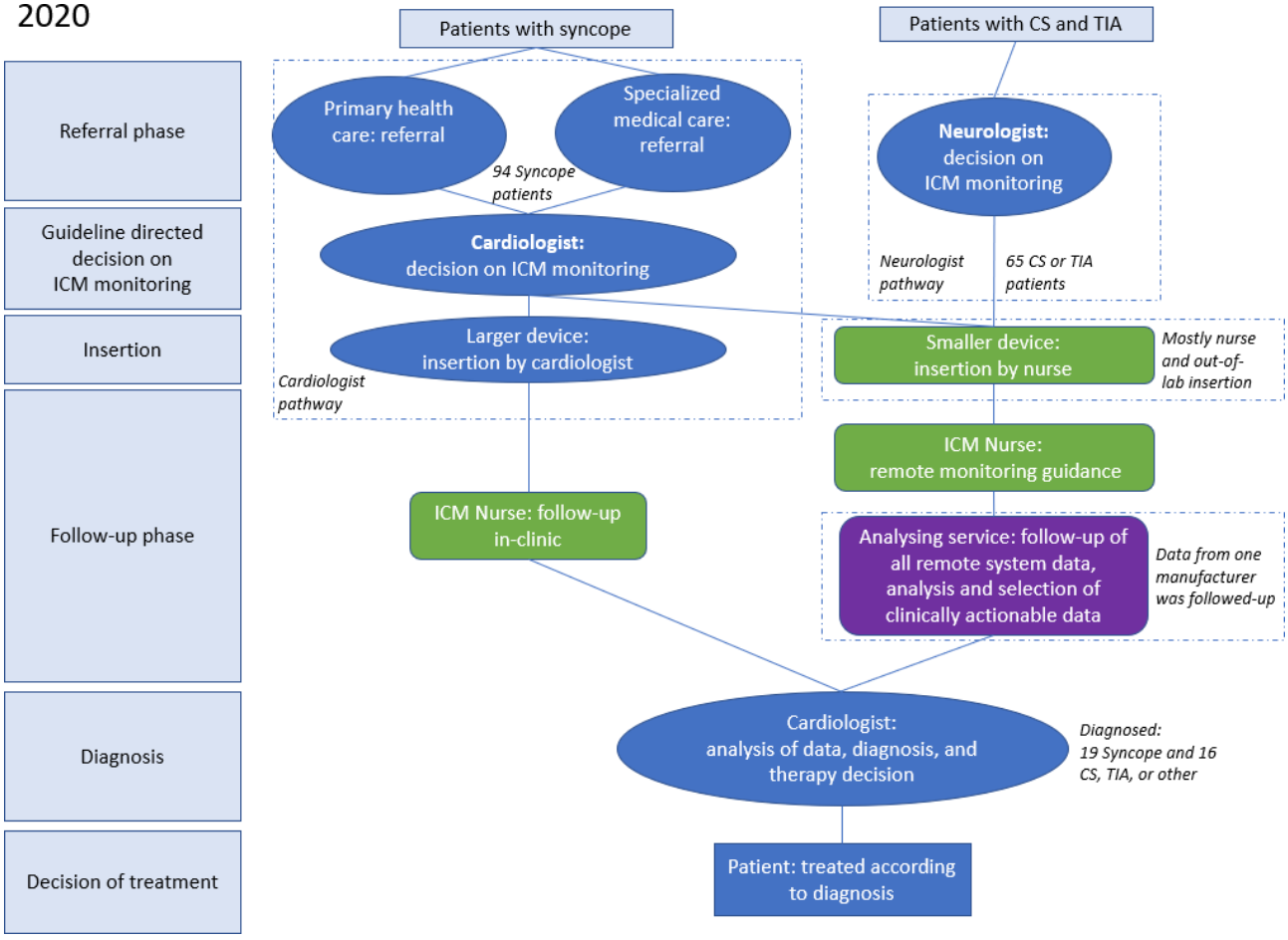



## Changes in the ICM Pathway as of 2020

### *Increasing Access to ICM Monitoring for Patients Who Had CS or TIA*

Based on the analysis, the clinical pathway was changed to improve its efficiency. The referral via cardiologist was a barrier for ICM monitoring for patients who had CS or TIA. To increase the access of patients who had CS, the neurologist could refer patients directly to an ICM procedure (Figure 2). Therefore, the decision on ICM insertions was transferred to the neurologist. This was in line with the updated 2020 ESC guidelines for AF management which had a stronger recommendation for ICM insertions for patients who had CS.

**Figure 2.** Patient pathways in 2020. CS: cryptogenic stroke; ICM: insertable cardiac monitor; TIA: transient ischemic attack.




### *Increasing Patients' Access to ICM Insertion Through Nurse-Inserted ICM in the Procedure Room*

The initial change focused on solutions for increasing the ICM insertion capacity of the hospital as well as patients' access to diagnostic services. Drawing from experiences abroad [6,9,14], where nurses safely and effectively conducted ICM insertions, the conclusion was made that training nurses to perform ICM procedures was safe and feasible.

The first ICM nurse-led insertion training program was initiated in Finland in 2019. The content was designed corresponding to the international, "nonphysician insert" ICM training program [6]. On the organizational level, the trained specialized nurses were deemed comparable to advanced practice providers as defined in international literature and publications [9]. Registered nurses underwent specialized training to perform ICM insertions (Multimedia Appendix 1). Based on the training and monitoring of 5 patients' ICM insertions under the supervision of a cardiologist, the Tampere Heart Hospital authorized 3 nurses to perform independent ICM insertions, thus officially delegating some of the physicians' responsibilities to the nurses officially to redistribute the workload.

Limited availability of the catheterization laboratory and management of the patient who had ICM workflow in the hospital led to launching nurse-led ICM insertions in a clean follow-up room specifically equipped for this procedure. The improved ICM clinical pathway with nurses performing ICM insertion of smaller devices was launched in the beginning of 2020. Larger ICMs were still on the market as well and cardiologists implanted them (Figure 2).

### *Outsourcing ICM Data Monitoring and Triaging*

Another notable change pertained to managing the workload associated with ICM data, as most ICMs were monitored remotely. Considering that a significant portion of the data were not clinically actionable and given the limitations in staff time, it was decided to outsource the first line analysis and triaging of remote follow-up data (Figure 2). The external monitoring service (FocusOn, Medtronic), consisting of technicians and rhythm cardiology professionals, analyzed the electrocardiogram data from patients who had ICM. They determined the urgency of the information and conveyed it to the hospital. This approach enables efficient data management, allowing hospital staff to focus on patients needing immediate attention [15] or perform additional ICM insertions.

## Methods

### Efficiency Assessment

A retrospective registry study was performed to assess the impact of the pathway changes. We computed key efficiency and safety metrics for the Tampere Heart Hospital before (2018) and after (2020) the change in the clinical pathways. Efficiency metrics included the number of patients treated with ICMs for unexplained syncope and CS or unexplained TIA, the number

of ICM insertions performed by nurses and cardiologists, procedure time, the number of insertions carried out in the catheterization laboratory, waiting time, diagnostic yield, and time to diagnosis. Clinically significant arrhythmia (bradycardia or tachycardia) was included in the diagnostic yield for patients who had syncope. For patients who had stroke, the diagnostic yield was measured as the proportion of patients with AF >6 minutes. Safety measures included the number of infections.

### Patient Population and Data Collection

Data collection encompassed all consecutive patients who had ICM at the Tampere Heart Hospital, irrespective of their indications, in the years 2018 and 2020. The data collection process was established as part of the clinic's ongoing medical care quality improvement efforts. Data were retrospectively collected from the patient records and procedure registry and identified using procedure codes and device serial numbers.

### Ethical Considerations

This study followed the ethical principles of the Declaration of Helsinki. Tampere University Hospital's Research Services of the Wellbeing services county of Pirkanmaa provided the permissions for the patient-level data collection from the electronic health record (R23641X). Because patients weren't contacted directly, informed consent wasn't required according to Finnish law. To protect patient privacy, patients who had ICM-level data were pseudonymized and subsequently aggregated into an anonymized format to prevent the identification of individuals. The data were handled according to the General Data Protection Regulation policy of the European Union.

### Statistical Analysis

Descriptive tabling of the quantitative variables was performed in Excel (version 2302; Microsoft 365 apps for enterprise). For categorical variables, the chi-square test was used to compare the distributions of 2 or more groups. For continuous variables, a 2-tailed *t* test was conducted to test for statistically significant differences. All calculations were carried out according to the intention to treat principle.

## Results

### Participants

In 2018, 74 consecutive patients were included in this study and in 2020, it was 159.

The proportion of female patients was 43.2% (n=32) and 51.6% (n=82) in 2018 and 2020, respectively. As they were being treated in an adult cardiology department, all patients were over 16 years of age. Most of the patients were aged between 40 and 79 years (n=58, 78.3%) in 2018, with a similar age distribution in 2020 (n=114, 71.7%). The median age of the patients was 66 (55.5-76.8) years in the 2018 patient population and 67 (54.0-75.0) years in the 2020 population. Participants' characteristics are presented in Table 1.

**Table .** Characteristics of participants who received ICM[a] insertions in 2018 and in 2020.

| | | 2018 (n=74), n (%) | 2020 (n=159), n (%) | *P* value |
|---|---|---|---|---|
| Sex (female) | | 32 (43.2) | 82 (51.6) | .24 |
| Age (years) | | | | .35 |
| | 16 - 39 | 4 (5.4) | 24 (15.1) | |
| | 40 - 59 | 22 (29.7) | 33 (20.8) | |
| | 60 - 79 | 36 (48.6) | 81 (50.9) | |
| | 80+ | 12 (16.2) | 21 (13.2) | |

[a]ICM: insertable cardiac monitor.

### Use of ICM According to Guidelines

In 2018, the indication for ICM insertion was mainly unexplained syncope (n=71, 95.9%) with 2.7% (n=2) of the patients indicated with CS. In contrast, in 2020, a total of 59.1% (n=94) were indicated with unexplained syncope and 39% (n=62) with CS. The number of patients receiving ICMs increased substantially from 2018 to 2020 ($P<.001$). For patients who had syncope, the increase was from 71 to 94. Notably, the use of ICMs in patients with SC or TIA substantially increased from 2018 (2 patients) to 2020 (62 patients; Table 2).

**Table .** Results—change in clinical pathway and safety.

| | | 2018 | 2020 | *P* value |
|---|---|---|---|---|
| Indication, n (%) | | | | <.001 |
| | Indication syncope | 71 (95.9) | 94 (59.1) | |
| | Indication cryptogenic stroke or TIA[a] | 2 (2.7) | 62 (39) | |
| | Other | 1 (1.4) | 3 (1.9) | |
| Waiting time to procedure (day), mean (SD) | | 49 (37.3) | 36 (27.7) | .02 |
| Nurse insertions, n (%) | | 0 (0) | 112 (70.4) | <.001 |
| Scheduled procedure time (min), n | | 60 | 45 | |
| Insertion in catheterization laboratory, n (%) | | 14 (18.9) | 3 (1.9) | <.001 |
| Overall complication rate, n (%) | | 2 (2.7) | 5 (3.1) | .85 |
| Data burden, n (%) | | | | <.001 |
| | Patients on remote monitoring | 38 (51.3) | 108 (67.9) | |
| | Patients on analyzing service | 0 (0) | 108 (67.9) | |

[a]TIA: transient ischemic attack.

## Waiting Time

A 2-sample *t* test was performed to compare the average waiting time from referral to insertion in 2018 and 2020. The average waiting time decreased significantly from 49 days in 2018 to 36 days in 2020 (*P*=.02; Table 2).

## Resource Use

In 2018, physicians conducted all insertions, while in 2020, 70.4% (n=112) of the ICM insertions were performed by specially trained nurses. The number of inserted ICMs doubled from 74 in 2018 to 159 in 2020. Delegating the responsibility of ICM insertions to trained nurses allowed physicians to allocate their time to other essential procedures and interventions. This transition to nurse-performed insertions in 2020 resulted in a saving of 48 hours (more than 6 working days) of physicians' time, a noteworthy improvement from the process in 2018 (Table 2).

## Catheterization Laboratory Use

In 2018, 18.9% (n=14) of the insertions were completed in the catheterization laboratory, whereas in 2020, this figure was reduced to 1.9% (n=3; *P*<.001). Additionally, the scheduled procedure time for ICM insertion decreased from 60 minutes in 2018 to 45 minutes in 2020. The streamlined procedure scheduling saved an additional 40 hours (1 wk) of nursing time and released the catheterization laboratory for other critical procedures, amounting to 27 hours per year (Table 2).

## Safety and Quality of the Procedure

All procedure-related complications were collected. The procedure-related complications were pain (1 patient in 2020), infection (2 patients in 2020), bleeding (2 patients in 2020), and device migration (1 patient in 2020). A total of 4 ICMs were explanted due to complications (3 relating to infection and 1 relating to pain). The complication rate remained consistent, with no significant change, at 2.7% (n=2) in 2018 and 3.1% (n=5) in 2020 (*P*=.85).

R-wave sensing data were only registered in 2020 after the initiation of nurse insertions. The average R-wave at implant in 2020 was 0.57 (SD 0.3) mV with 8 (5%) patients having an R-wave below 0.2 mV.

## Nurse Productivity

Remote monitoring was set up for 51.3% (n=38) of the patients in 2018 and for 67.9% (n=108) in 2020. In 2018, none of the remote-monitored patients who had ICM were followed up by an outsourced analyzing service, while in 2020, all ICM remote-monitored patients (n=108) were in the FocusOn-system. In 2018, nurses were responsible for analyzing all remote transmissions, consuming a substantial amount of their time. The number of transmissions that needed analyzing from nurses was not available. In 2020, the initial review and triaging of remote transmissions were outsourced to an external monitoring center. This external service escalated 11.2% (204 out of 1817) of the transmissions to the clinic for review. Assuming an average of 11 minutes per transmission by a nurse [8,10,16], this external service saved 296 hours (approximately 40 working days corresponding to almost 2 mo) of nursing time in 2020 (Table 2).

## Diagnostic Yield

Notably, the quality of the diagnostic pathway was high, with a high diagnostic yield despite the increase in inserted ICMs from 2018 to 2020 (Table 3). The 1-year diagnostic yield for patients with syncope remained high and exhibited no statistically significant difference between 2018 and 2020 (n=19, 26.7% vs n=19, 20.2%; *P*=.32). The 36-month diagnostic yield for patients who had syncope was generally high, with no statistically significant difference between 2020 (n=36, 38.3%) and 2018 (n=32, 45.1%; *P*=.38). The time to diagnosis was not

statistically significantly different in 2018 and 2020 for patients who had syncope (109 vs 114 days; *P*=.88). Further information of detected arrhythmias is included in Multimedia Appendix 2.

**Table .** Diagnostic yield-intention to treat (2018: n=74; 2020: n=159).

| | 12 month follow-up, n (%) | | *P* value | 24 month follow-up, n (%) | | *P* value | 36 month follow-up, n (%) | | *P* value |
|---|---|---|---|---|---|---|---|---|---|
| | 2018 | 2020 | | 2018 | 2020 | | 2018 | 2020 | |
| Overall | 19 (25.7) | 35 (22) | .54 | 31 (41.9) | 52 (32.7) | .17 | 32 (43.2) | 63 (39.6) | .60 |
| Syncope | 19 (26.7) | 19 (20.2) | .32 | 31 (43.7) | 31 (33) | .16 | 32 (45.1) | 36 (38.3) | .38 |
| Stroke | 0 (0) | 17 (27.4) | N/A[a] | 0 (0) | 21 (33.9) | N/A | 0 (0) | 27 (43.5) | N/A |

[a]N/A: not applicable.

The 1-year diagnostic yield (AF diagnosis) for patients who had CS was 27.4% (n=17) and the 36-month diagnostic yield was 43.5% (n=27) in 2020. The average time to diagnosis for patients who had stroke was 127 days in 2020.

## Discussion

### Principal Findings

Our study illustrated that the shift from physician-led ICM insertions to a clinical pathway where nurses inserted the majority of ICMs released a substantial amount of staff time and resources without compromising the quality of the clinical pathway. The efficiency assessment showed that nurse insertion and the use of an external monitoring and triaging service significantly improved the use of hospital resources, such as patient access to ICM insertion, follow-up, and diagnosis. The results correspond to findings from the UK's National Health Service health care system, where trained nurses have independently been taking care of ICM insertions and follow ups with high quality treatment and safety since 2015 [6].

Regarding the patient follow-up, while in 2018 nurses analyzed all remote monitoring data, in 2020 that part of the workflow was outsourced to an external monitoring and triaging service. As nurses in 2020 monitored only those remote transmissions that were escalated, they could perform more ICM insertions and actionable patient follow-ups. Similar efficiency benefits from outsourcing part of the workflow have been reported previously [10,17]. According to Giannola et al [17], the introduction of such service offered efficiency and effectiveness in patient care more safely than when compared with remote follow-up handled solely at hospital level. Outsourcing the management of remote monitoring data has been seen as a key tool for saving staff time [8,18]. In addition, Biundo et al [8] highlighted the need for appropriate staff resources to support patient management activities, including remote monitoring. Considering the heterogeneity in the infrastructure and staff capacity of hospitals managing patients who had ICM, different organizational models should be considered locally to achieve efficient patient management, including outsourcing part of the remote monitoring workflow [15]. Although the use of an outsourced triaging service will add some costs, more efficient use of hospitals resources and increased number of insertions will probably help hospitals to reclaim the costs from the health care funding system.

Our study at the Tampere Heart Hospital showed both a decrease in the waiting time for the procedure and an increase in the number of patients receiving care in response to the implemented changes. Overall, the number of ICM insertions in 2020 doubled, with indications for CS and TIA also increasing significantly from 2018 to 2020.

The new workflow enabled nurses to gain new skills and broader responsibilities, while physicians could refocus on specialized care. Additionally, the shorter procedure released overall staff time in 2020 compared with 2018. In this study, we only had access to scheduled procedure time and not the actual procedure time. However, these results correspond to the findings of Lim et al [6] with the study conducted in the National Health Service.

In addition, the Tampere Heart Hospital catheterization laboratory was released for other procedures, as the insertions performed in this setting decreased significantly. Rogers et al [16] showed similar results for insertions performed outside the catheterization laboratory. Moving the procedure to office settings saved time spent by patients in hospital, space and resources used, clinical staff time, and, thus, the total costs of the procedure [16]. When aiming to increase efficiency in the clinical pathway, a detailed analysis of all resources supports optimizing the process.

In this study, only cardiac arrhythmia diagnoses were included in the reporting of the diagnostic yield. Furthermore, an "intention to treat" principle was used, hence all patients were included with full follow-up time, even though they were diagnosed, deceased, or exited the population earlier for any other reason.

In our study, the diagnostic yields for patients who had syncope were high both in 2018 and 2020 (n=32, 45.1% and n=36, 38.3%; *P*=.38). In a meta-analysis by Solbiati et al [18], the overall diagnostic yield was reported to be similar to our study (43.9%) [18].

Sanna et al [19] reported the AF detection rate for patients who had stroke to be 12.4% at the 12-month follow-up and 30% at the 36-month follow-up [19]. Our study showed an even higher diagnostic yield of 43.5% (n=27) at 36 months. Notably, the patient population in the initial care pathway only included a very low number of patients who had CS or TIA which prevents a comparison between 2018 and 2020 for this indication [19]. As almost half of the patients who had syncope and patients who had stroke receive a cardiac arrhythmia diagnosis after ICM insertion, there could be underuse of ICMs in both patient

groups. There is also a risk for overdiagnosing patients with clinically insignificant arrhythmias and this leading to a potentially harmful therapy (eg, pacemaker implantation after asymptomatic night-time bradyarrhythmia or anticoagulating patient with very short device-detected AF). Choosing patients for ICM insertion is a demanding task and choosing a therapy after device-detected arrhythmia is even more complex. Further studies are needed to address these problems.

Importantly, the changes in the ICM pathway did not compromise patient safety. In this study, the complication rate did not change significantly regardless of whether the procedure was performed solely by a physician in the catheterization laboratory or a procedure room (n=2, 2.7%) or mainly by a nurse in a procedure room (n=5, 3.1%). As the sample size of our study is quite small, even 1 complication will have a significant impact on reported percentages. In earlier studies, procedure-related adverse events have been between 1.1% and 2.6% depending on the location of the procedure [20,21], and the complication rate has been 1% for nurse-performed ICM insertions and 2.2% for physician-performed insertions [6].

At the time of launching this study, there was only 1 other hospital in Finland that had initiated nurse-led insertions. At the time of publishing these results, Finland had 9 hospitals running nurse-led ICM processes. A prospective study assessing the cost-effectiveness of a nurse-led ICM process more precisely could lead to implementing these changes in other health care systems as well.

### Limitations

This study has several limitations. First, it is a single center study with a small number of consecutive patients who had ICM without randomization. Nonetheless, they represent patients from a tertiary level cardiac hospital that serves a population of 520,000 inhabitants [22]. The real-world setting helps to describe how a clinical pathway change is made in practice. Second, the retrospective analysis uses data that was documented or available in the electronic health record. For example, the working time that the nurses used to analyze the data for the 74 patients was not recorded at that time. Therefore, for the efficiency estimation concerning the saved working time of nurses, we used only the 2020 data in comparison with earlier research. Third, R-waves were only measured after the workflow shift to nurse insertions. However, the measured R-wave amplitudes are in line with previously published results [23].

### Conclusions

The change in the clinical pathway to nurse-perfomed insertion in a procedure room and the use of an external monitoring and triaging service significantly improved the efficiency of the pathway for patients indicated for an ICM. In addition, nurse-led insertion released a significant amount of staff time and resources without compromising the quality of the treatment. It can be stated that clinical pathway improvements enable offering ICMs to a greater number of patients to meet the diagnostic demand.

## Acknowledgments

The authors of this paper would like to thank the Tampere Heart Hospital team for permitting the observation of their insertable cardiac monitor workflow and participation in the data collection. Data analysis was performed by Medtronic. This research did not receive a specific grant from any funding agency in the public, commercial, or not-for-profit sectors. This paper was proofread by Merja Kalima, MA, from Jamk University of Applied Sciences.

## Data Availability

The datasets generated or analyzed during this study are not publicly available due to the European Union's General Data Protection Regulation regulations but are available from the corresponding author on reasonable request in anonymized form.

## Authors' Contributions

VV, VM, PK, and OS handled the change in pathway. VM, OS, and MLJ collected the data. OS worked on this study's design and the writing of the first draft of this paper. MLJ and OS analyzed the anonymized data. VV, PK, OS, JH, MLJ, JV, and EN revised this paper. All authors reviewed and contributed to the final paper.

## Conflicts of Interest

OS, MLJ, JV, and EN are Medtronic employees and shareholders. Medtronic paid the submission fee.

Multimedia Appendix 1
ICM nurse insertion training program. ICM: insertable cardiac monitor.
[DOCX File, 19 KB - cardio_v9i1e67774_app1.docx ]

Multimedia Appendix 2
Arrhythmias detected.
[DOCX File, 17 KB - cardio_v9i1e67774_app2.docx ]

## References


1. Brignole M, Moya A, de Lange FJ, et al. 2018 ESC Guidelines for the diagnosis and management of syncope. Eur Heart J 2018 Jun 1;39(21):1883-1948. [doi: 10.1093/eurheartj/ehy037] [Medline: 29562304]
2. Cheung CC, Krahn AD. Loop recorders for syncope evaluation: what is the evidence? Expert Rev Med Devices 2016 Nov;13(11):1021-1027. [doi: 10.1080/17434440.2016.1243463]
3. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016 Oct 7;37(38):2893-2962. [doi: 10.1093/eurheartj/ehw210]
4. Dulai R, Hunt J, Veasey RA, Biyanwila C, O'Neill B, Patel N. Immediate implantable loop recorder implantation for detecting atrial fibrillation in cryptogenic stroke. J Stroke Cerebrovasc Dis 2023 Mar;32(3):106988. [doi: 10.1016/j.jstrokecerebrovasdis.2023.106988]
5. Ferrick AM, Raj SR, Deneke T, et al. 2023 HRS/EHRA/APHRS/LAHRS expert consensus statement on practical management of the remote device clinic. Heart Rhythm 2023 Sep;20(9):e92-e144. [doi: 10.1016/j.hrthm.2023.03.1525]
6. Lim WY, Papageorgiou N, Sukumar SM, et al. A nurse‐led implantable loop recorder service is safe and cost effective. J Cardiovasc Electrophysiol 2019 Dec;30(12):2900-2906. [doi: 10.1111/jce.14206]
7. Timmis A, Vardas P, Townsend N, et al. European Society of Cardiology: cardiovascular disease statistics 2021: executive summary. Eur Heart J Qual Care Clin Outcomes 2022 Jun 6;8(4):377-382. [doi: 10.1093/ehjqcco/qcac014]
8. Biundo E, Burke A, Rosemas S, Lanctin D, Nicolle E. Clinic time required to manage cardiac implantable electronic device patients: a time and motion workflow evaluation. Eur Heart J 2020 Nov 1;41(Supplement_2):ehaa946.0821. [doi: 10.1093/ehjci/ehaa946.0821]
9. Kipp R, Young N, Barnett A, et al. Injectable loop recorder implantation in an ambulatory setting by advanced practice providers: analysis of outcomes. Pacing Clinical Electrophis 2017 Sep;40(9):982-985 [FREE Full text] [doi: 10.1111/pace.13155]
10. Nicolle E, Lanctin D, Rosemas S, De Melis M. Clinic time required to manage remote monitoring of cardiac implantable electronic devices: impact of outsourcing initial data review and triage. EP Europace 2021 May 24;23(Supplement_3):euab116.519. [doi: 10.1093/europace/euab116.519]
11. Sane M, Annukka M, Toni J, et al. Real‐life data on the workload of cardiac implantable electronic device remote monitoring in a large tertiary center. Pacing Clinical Electrophis 2023 Sep;46(9):1109-1115. [doi: 10.1111/pace.14792]
12. Kuntien työvoimaennuste 2030: hoitajissa, sosiaalityöntekijöissä, ja lastentarhanopettajissa suurin osaajapula nyt ja tulevaisuudessa [Web page in Finnish]. KEVA. 2021. URL: https://www.keva.fi/uutiset-ja-artikkelit/kuntien-tyovoimaennuste-2030-hoitajissa-sosiaalityontekijoissa-ja-lastentarhanopettajissa-suurin-osaajapula-nyt-ja-tulevaisuudessa/ [accessed 2025-03-05]
13. Kirkonpelto TM, Mäntyranta T. Toimeenpanosuunnitelma 2024–2027: Sosiaali- ja terveysalan sekä pelastusalan henkilöstön riittävyyden ja saatavuuden turvaaminen [Article in Finnish]. Sosiaali- ja terveysministeriön julk 2024;11:1-38 [FREE Full text]
14. Eftekhari H, He H, Lee JD, et al. Safety and outcome of nurse-led syncope clinics and implantable loop recorder implants. Heart Rhythm 2022 Mar;19(3):443-447. [doi: 10.1016/j.hrthm.2021.11.006] [Medline: 34767989]
15. Seiler A, Biundo E, Di Bacco M, et al. Clinic time required for remote and in-person management of patients with cardiac devices: time and motion workflow evaluation. JMIR Cardio 2021 Oct 15;5(2):e27720. [doi: 10.2196/27720] [Medline: 34156344]
16. Rogers JD, Piorkowski C, Sohail MR, et al. Resource utilization associated with hospital and office-based insertion of a miniaturized insertable cardiac monitor: results from the RIO 2 randomized US study. J Med Econ 2020 Jul 2;23(7):706-713. [doi: 10.1080/13696998.2020.1746548]
17. Giannola G, Torcivia R, Farulla RA, Cipolla T. Outsourcing the remote management of cardiac implantable electronic devices: medical care quality improvement project. JMIR Cardio 2019 Dec 18;3(2):e9815. [doi: 10.2196/cardio.9815] [Medline: 31845898]
18. Solbiati M, Casazza G, Dipaola F, et al. The diagnostic yield of implantable loop recorders in unexplained syncope: a systematic review and meta-analysis. Int J Cardiol 2017 Mar 15;231(231):170-176. [doi: 10.1016/j.ijcard.2016.12.128] [Medline: 28052814]
19. Sanna T, Diener HC, Passman RS, et al. Cryptogenic stroke and underlying atrial fibrillation. N Engl J Med 2014 Jun 26;370(26):2478-2486. [doi: 10.1056/NEJMoa1313600] [Medline: 24963567]
20. Mittal S, Sanders P, Pokushalov E, et al. Safety profile of a miniaturized insertable cardiac monitor: results from two prospective trials. Pacing Clin Electrophysiol 2015 Dec;38(12):1464-1469. [doi: 10.1111/pace.12752] [Medline: 26412309]
21. Wong GR, Lau DH, Middeldorp ME, et al. Feasibility and safety of Reveal LINQ insertion in a sterile procedure room versus electrophysiology laboratory. Int J Cardiol 2016 Nov;223(223):13-17. [doi: 10.1016/j.ijcard.2016.08.113]
22. Pirkanmaan hyvinvointialue valmiussuunnitelman yleinen osa [Web page in Finnish]. Pirkanmaan Hyvinvointialue. 2022. URL: https://pirha.cloudnc.fi/download/noname/%7B387a13c5-faec-47c5-8e38-03d36ecf59ab%7D/39712 [accessed 2025-02-08]

23. Pürerfellner H, Sanders P, Pokushalov E, et al. Miniaturized Reveal LINQ insertable cardiac monitoring system: first-in-human experience. Heart Rhythm 2015 Jun;12(6):1113-1119. [doi: 10.1016/j.hrthm.2015.02.030] [Medline: 25728756]

## Abbreviations

**AF:** atrial fibrillation
**CS:** cryptogenic stroke
**ESC:** European Society of Cardiology
**ICM:** insertable cardiac monitor
**TIA:** transient ischemic attack

*Edited by A Coristine; submitted 21.10.24; peer-reviewed by C Monkhouse, M Richards; revised version received 18.12.24; accepted 27.12.24; published 21.03.25.*

*Please cite as:*
*Vanhala V, Surakka O, Multisilta V, Lundsby Johansen M, Villinger J, Nicolle E, Heikkilä J, Korhonen P*
*Efficiency Improvement of the Clinical Pathway in Cardiac Monitor Insertion and Follow-Up: Retrospective Analysis*
*JMIR Cardio 2025;9:e67774*
*URL: https://cardio.jmir.org/2025/1/e67774*
*doi:10.2196/67774*

Original Paper

# Telehealth Support From Cardiologists to Primary Care Physicians in Heart Failure Treatment: Mixed Methods Feasibility Study of the Brazilian Heart Insufficiency With Telemedicine Trial

Leonardo Graever[1,2], MSc, MD; Priscila Cordeiro Mafra[3], MSc, MD; Vinicius Klein Figueira[4], MD; Vanessa Navega Miler[4,5], MS; Júlia dos Santos Lima Sobreiro[4,6], MS; Gabriel Pesce de Castro da Silva[4], MD, MSc; Aurora Felice Castro Issa[4,6], MD, PhD; Leonardo Cançado Monteiro Savassi[7], MD, PhD; Mariana Borges Dias[8], MD; Marcelo Machado Melo[4], MD; Viviane Belidio Pinheiro da Fonseca[4], MD; Isabel Cristina Pacheco da Nóbrega[4], BSc, MSc; Maria Kátia Gomes[2,3], MD, PhD; Laís Pimenta Ribeiro dos Santos[9], BSN, MSc; José Roberto Lapa e Silva[2], MD, PhD; Anne Froelich[10], MD, PhD; Helena Dominguez[1,11], MD, PhD

[1]Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[2]Departamento de Clínica Médica, Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[3]Instituto de Atenção à Saúde São Francisco de Assis, Centro de Ciências da Saúde, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[4]Instituto Nacional de Cardiologia, Rio de Janeiro, Brazil

[5]Faculdade de Medicina, Universidade Federal Fluminense, Niterói, Brazil

[6]Faculdade de Medicina, Instituto de Educação Médica, Rio de Janeiro, Brazil

[7]Departamento de Medicina de Família e Comunidade, Saúde Mental e Coletiva, Escola de Medicina, Universidade Federal de Ouro Preto, Ouro Preto, Brazil

[8]Ministério da Saúde, Brasília, Brazil

[9]Escola Nacional de Saúde Pública, FIOCRUZ, Rio de Janeiro, Brazil

[10]Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[11]Department of Cardiology, Bispebjerg Hospital, Copenhagen, Denmark

**Corresponding Author:**
Helena Dominguez, MD, PhD
Department of Cardiology
Bispebjerg Hospital
Bispebjerg Bakke 23
Copenhagen, 2400
Denmark
Phone: 45 22 98 93 43
Email: mdom0002@regionh.dk

## Abstract

**Background:** Heart failure is a prevalent condition ideally managed through collaboration between health care sectors. Telehealth between cardiologists and primary care physicians is a strategy to improve the quality of care for patients with heart failure. Still, the effectiveness of this approach on patient-relevant outcomes needs to be determined.

**Objective:** This study aimed to assess the feasibility of telehealth support provided by cardiologists for treating patients with heart failure to primary care physicians from public primary care practices in Rio de Janeiro, Brazil.

**Methods:** We used mixed methods to assess the feasibility of telehealth support. From 2020 to 2022, we tested 2 telehealth approaches: synchronous videoconferences (phase A) and interaction through an asynchronous web platform (phase B). The primary outcome was feasibility. Exploratory outcomes were telehealth acceptability of patients, primary care physicians, and cardiologists; the patients' clinical status; and prescription practices. Qualitative methods comprised content analysis of 3 focus groups and 15 individual interviews with patients, primary care physicians, and cardiologists. Quantitative methods included the baseline assessment of 83 patients; a single-arm, before-and-after assessment of clinical status in 58 patients; and an assessment of guideline-directed medical therapy in 28 patients with reduced ejection fraction measured within 1 year of follow-up. We

integrated qualitative and quantitative data using a joint display table and used the A Process for Decision-Making After Pilot and Feasibility Trials framework for feasibility assessment.

**Results:** Telehealth support from cardiologists to primary care physicians was generally well accepted. As barriers, patients expressed concern about reduced direct access to cardiologists, primary care physicians reported work overload and a lack of relative advantage, and cardiologists expressed concern about the sustainability of the intervention. Quantitative analysis revealed an overall poor baseline clinical status of patients with heart failure, with 53% (44/83) decompensated, as expected. Compliance with guideline-directed medical therapy for the treatment of heart failure with reduced ejection fraction after telehealth showed a modest improvement for β-blockers (17/20, 85% to 18/19, 95%) and renin-angiotensin-aldosterone system inhibitors (14/20, 70% to 15/19, 79%) but a drop in the prescription of spironolactone (16/20, 80% to 15/20, 75%). Neprilysin and sodium-glucose cotransporter 2 inhibitors were introduced in 4 and 1 patient, respectively. Missing record data precluded a more precise analysis. The feasibility assessment was positive, favoring the asynchronous modality. Potential modifications include more effective patient and professional recruitment strategies and educational activities to raise awareness of collaborative support in primary care.

**Conclusions:** Telehealth was feasible to implement. Considering the stakeholders' views and insights on the process is paramount to attaining engagement. Missing data must be anticipated for future research in this setting. Considering the recommended adaptations, the intervention can be studied in a cluster-randomized trial.

***(JMIR Cardio 2025;9:e64438)*** doi:10.2196/64438

**KEYWORDS**

heart failure; telemedicine; telehealth; intersectoral collaboration; primary health care; low- and middle-income countries; family practice

## Introduction

### Background

Collaboration among health care professionals is essential for delivering the best possible care for the population [1]. Telehealth, defined in this paper as the interaction between health care professionals using remote communication tools to collaborate on patient care [2,3], may increase the efficiency of health care systems, reduce costs, and improve patients' quality of life while lowering the need for in-person appointments with specialists and referrals [4,5]. Specifically, chronic disease management involving multidisciplinary collaboration is known to improve the quality of care [6,7].

Heart failure is a chronic condition and the end stage of many cardiovascular diseases, with a significant impact on public health [8-10]. Recent epidemiologic studies on the global burden of disease point to an incidence of up to 20 cases per 1000 persons per year and a prevalence of 1% to 3% of the population, affecting 64 million people worldwide [11-14]. Readmission rates can be as high as 40% in 6 months [15], burdening health systems with an estimated annual cost of US $108 billion worldwide [16]. The 5-year specific mortality rate may reach 75%, and quality of life is jeopardized. Population aging, the increase in survival rates after acute cardiologic events, and better access to health care will increase the prevalence of heart failure by up to 8.5% in 2030 according to prediction models [17].

Notwithstanding the unfavorable epidemiological scenario, heart failure is amenable to pharmacological treatment and behavior change. Most interventions can be delivered in primary care [18,19] and other outpatient settings with positive results [20,21], and new guidelines, including novel pharmacological options, are published and updated frequently [22,23]. Nevertheless, the overall physician adherence to the recommendations is low. The proportion of patients with heart failure with reduced ejection fraction (HFrEF) treated following guideline-directed medical therapy (GDMT) is reported as 27% to 73%, constituting only 14% when reaching target doses is considered [24]. Primary care physicians with a general medicine background commonly need support in assisting these patients, as described in previous studies [25-28]. Therefore, there is plenty of room for improvement, making it a suitable case for collaborative strategies such as telehealth.

Telehealth services have been commonly used as a collaborative care strategy, mainly in North America and, to a lesser extent, in Europe [29], with positive results [30,31]. They are less common in low- and middle-income countries. Brazil has a national telehealth program named *Telessaúde Brasil Redes* [32], which aims to foster the development of telehealth nuclei in Brazilian states and regions. At least 3 large telehealth services have been implemented in the last decades. Unfortunately, reports about telehealth implementation in Brazil have pointed to low adoption rates by primary care physicians [33-37].

Implementation research studies indicate that telehealth implementation, as a complex intervention, is influenced by multiple factors that may facilitate or undermine its adoption and usability [38-40]. Telehealth adoption is below the expected level in many settings due to subjective factors such as resistance to innovation and practical aspects such as infrastructure availability, technical challenges, communication hardships between sectors, and work overload from other tasks [41-43]. Furthermore, solid, high-quality evidence of the benefit of telehealth, especially in assessing patient-relevant outcomes, is lacking [44]. Recently published systematic reviews point to the need for trials with enough statistical power focusing on patient-relevant outcomes such as mortality, hospital admissions, and quality of life [4,29,44,45]. For all the reasons and knowledge gaps described previously, we designed a clinical

trial [46] within the Brazilian Heart Insufficiency With Telemedicine (BRAHIT) frame project, an academic collaboration between medical researchers from Denmark and Brazil's higher education and health institutions [47]. The trial aims to evaluate whether telehealth support from cardiologists to primary care physicians improves the quality of heart failure management and impacts patient-relevant outcomes.

As recommended by most frameworks for studying complex interventions [48,49], we previously tested the implementation of the intervention used in this study, aiming to assess the feasibility of the telehealth process designed as the trial intervention. We tested a synchronous approach, where real-time case discussions are held between specialists and primary care physicians using remote communication tools (eg, videoconference), and an asynchronous approach, where the communication does not require real-time contact between the parties and the remote interaction happens using a non–real-time strategy (eg, SMS text messages).

We aimed to answer the following research question: is it feasible to implement telehealth support from cardiologists to primary care physicians in the clinical practice settings of Rio de Janeiro and evaluate it as an intervention within a cluster-randomized trial? Other pertinent research questions included the following: which factors influence primary care physicians' adoption of telehealth support? How do other stakeholders, such as patients and teleconsulting cardiologists, perceive the intervention? Does telehealth support alter current clinical practices among primary care physicians?

### Objectives

This study aimed to analyze factors influencing the delivery and acceptability of telehealth support by primary care physicians, cardiologists, and patients (stakeholders), including context factors, facilitators, barriers, opportunities, and threats, and analyze whether telehealth support influences primary care physicians' treatment practices and the clinical status of patients with heart failure.

## Methods

### Study Design

This was a prospective study using mixed methods and a concurrent design. The qualitative approach included thematic analysis of data from focus groups and individual interviews with the participants using predefined, semistructured scripts. The analysis followed an inductive, constructivist approach. We sought data about the context and the telehealth execution, drawing connections between our preconceived hypotheses and assumptions (theories) and the collected data guided by the content analysis methodology by Bardin [50]. We chose this design to collect and analyze descriptive and subjective in loco information that could help us answer our research questions. The quantitative assessment involved a descriptive analysis of the patients' clinical changes, including vital signs, symptoms, and prescribed medications in the cases discussed.

For reporting guidance, we used, where applicable, the CONSORT (Consolidated Standards of Reporting Trials) extension for pilot and feasibility trials [51], the Strengthening the Reporting of Observational Studies in Epidemiology statement for observational research [52], the Standards for Reporting Qualitative Research statement [53], the recommendations by Braun and Clarke [54] for reporting qualitative studies, guidelines for reporting mixed methods studies [55], and additional guiding literature [56,57].

### Setting

The BRAHIT project started in 2019 with the principal aim of implementing digital solutions to improve the quality of cardiovascular disease care in Rio de Janeiro, Brazil's second-largest city with 6.2 million inhabitants. Brazil's population relies on a universal health system with free access to comprehensive care, and Brazil has invested in primary care through the implementation of the Family Health Strategy over the last 25 years [58]. In this context, Rio de Janeiro has been the setting for significant primary care reforms in the previous 15 years, showing a marked increase in health care structure and workforce [59]. There are currently 238 primary health care practices in the city hosting 1352 teams, each composed of 1 physician, 1 nurse, 1 nurse technician, and 5 to 6 community health workers. Primary care practices also deliver oral health care and have the support of mental health and rehabilitation professionals.

As one of the main cities in the country and former capital, Rio de Janeiro also hosts a thorough specialized service network, including national institutes such as the National Institute of Cardiology (INC), whose team was responsible for the telehealth support to the primary care teams in this study. The choice of telehealth as the studied intervention within the BRAHIT project relied on the strategic role of collaborative interactions between health services to improve health care [6], which aligned with the project's main strategic goal.

Other BRAHIT project research activities include a systematic review of telehealth and a cluster-randomized trial registered at ClinicalTrials.gov (NCT04466852), which was in the recruitment phase when this paper was submitted.

### Intervention

#### *Overview*

The intervention assessed in this study was telehealth support requested by a primary care physician to discuss a heart failure case and executed by a cardiologist from the INC. The intervention aimed to support general physicians in dealing with the clinical aspects of heart failure management, including diagnostic, treatment, and referral practices. The feasibility study and interventions were organized in 2 different phases and approaches. Telehealth occurred through scheduled synchronous videoconferences or an asynchronous texting and data exchange platform depending on the study phase, as described in the following sections.

#### *Phase A: Synchronous Videoconferences*

Phase A started in August 2020, when videoconferences (synchronous approach) between cardiologists and primary care physicians were implemented to discuss cases of patients with heart failure from one of the Rio de Janeiro municipality's primary care practices. The practice comprised 15 primary care

teams. As one of the hosts of the family medicine residency program in Rio de Janeiro, it also has 2 family medicine residents per team (year 1 and year 2) in addition to the original team composition described previously. This practice provides primary care for >45,000 people in a socioeconomically deprived area.

The research team presented the BRAHIT project's telehealth support offer to a group of physicians from the practice who could disseminate the information to the remaining staff members and agreed on the methods. A web-based schedule was organized and hosted on the practice's Google workspace, where the primary care physicians could schedule the telehealth session with the cardiologists.

In a preliminary meeting, all participants were previously trained in telehealth by one of the researchers (LG). In total, 1 to 3 cases of patients with heart failure were discussed in each session, which could take place once a week unless there was no appointment. The primary care physicians used the practice's computers, and the cardiologists used the INC research department computers to connect and interact via the Zoom platform (Zoom Video Communications) licensed for the project. Phase A lasted from August 2020 to June 2021 (11 months).

### *Phase B: Asynchronous Telehealth Using an Online Platform*

Phase B started in July 2021, when the researchers decided to upscale the telehealth offer to all other primary care practices in the city. An IT company was hired to develop an online platform conceived by the researchers and based on similar experiences described in the literature [60] to allow for information exchange via text (asynchronous), substituting videoconferences as the initial interaction tool. The web-based platform was hosted on the project's website (Figure 1).

Upon registration and secure access granted by the research data management team (Figure 2), the primary care physicians entered their professional identification and contact information, the patient's demographic and clinical data, and the reason for telehealth.

The research group's teleconsultant should respond within 2 working days through a texting service within the platform. If primary care physicians deemed it necessary, they could still make synchronized phone or videoconference calls on demand. In this case, after agreeing with the cardiologist, they would use the WhatsApp app (Meta Platforms) for voice or video calls at their discretion. The web-based platform did not offer synchronous contact in the form of audio or video calls due to time and financial constraints for the tool's development.

One of the researchers (LG) shared the BRAHIT project's telehealth offer through presentations to the municipal health department, the regional primary care health coordination offices, and the family medicine residency program staff. In this second phase, 13 primary care practices participated in the telehealth program, including the practice involved in phase A. While primary care physicians could discuss cases of patients with other cardiologic diagnoses, this study focused solely on the discussion of heart failure cases.

In both phases, the duration of support was at the discretion of the primary care physicians. Regardless of the study phase, all patients had access to standard care, including consultations with physicians and nurses, preventive measures, oral health treatments, and follow-up visits from community agents. Participating primary care teams received weight scales, automatic blood pressure monitors, and oximeters to encourage patient follow-up. Phase B lasted from July 2021 to December 2022 (19 months).

**Figure 1.** Telehealth online platform landing page used in all study phases for intervention delivery (provider-to-provider support from cardiologists to primary care physicians via telehealth) from August 2020 to December 2022. Permission obtained by the authorship for the use of the image without attribution.

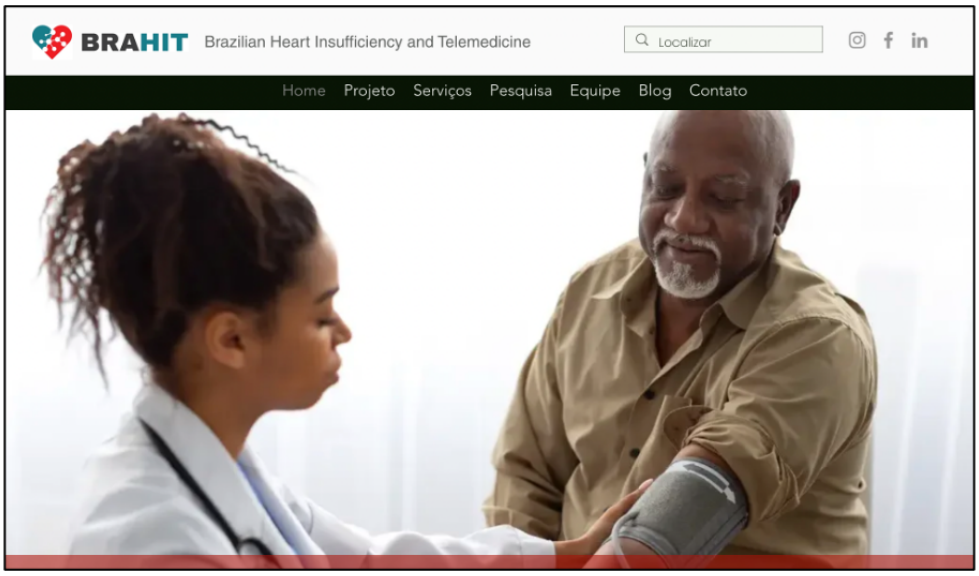
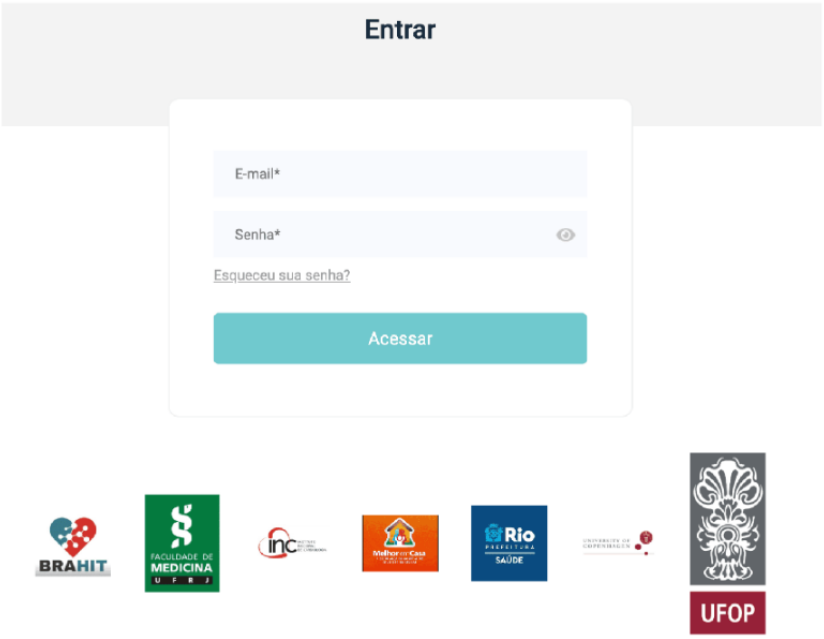


**Figure 2.** Log-in page for the online platform, restricted to registered users to protect data access and ensure their safety.




## Participants and Data Collection

### *Qualitative Methods*

We conducted 3 separate focus groups (group 1, group 2, and group 3) after the end of phase A and 15 interviews after phase B. The first author, LG, a physician and PhD candidate, scheduled, organized, and conducted the focus group sessions, whereas PCM, a female physician and master's degree candidate, conducted the individual interviews. Both are trained in executing qualitative research data collection. MKG, a female researcher with robust qualitative research experience, supervised and supported data collection and analysis.

At the beginning of all focus group sessions, LG explained the research and session objectives and disclaimed the research objectives and premises, including the group's assumptions and theories. Probing questions were used as an orientation for each focus group to facilitate the meeting interactions. All meetings were audio recorded for later transcription and content analysis. The probing questions of the semistructured interview script were about telehealth within the BRAHIT project, its use in the practices, and participants' perception of their ability to manage patients with heart failure.

For group 1, researchers MKG and LG invited all the primary care physicians from the phase A practice, including family and community medicine specialists or residents. Considering the initial response of 5 family physicians and 10 residents, the researchers decided to conduct 1 session because a second one could have low attendance due to the participants' time constraints. All invitees attended the session. With one exception, most participants were young physicians who had graduated in the previous 10 years. They are an engaged, proactive health care team that is usually cooperative and prone to quality improvement initiatives. All primary care physicians using telehealth and participating in this focus group were members of the Rio de Janeiro municipality's family medicine residency program. This could have contributed to better engagement and assessment of educational activities such as telehealth. One of the primary care physicians was assigned as the observer. The session, which lasted 96 minutes, took place on June 22, 2021, in the practice auditorium.

For group 2, all 5 cardiologists who provided telehealth support during the study were considered eligible for the session and invited. The cardiologists have a strong connection with the researchers and vice versa as they are also project workers or researchers. In total, 80% (4/5) of the invited cardiologists attended the focus group session. One could not be contacted and had already left the project team. The senior author (HD) participated as an observer. The age range of the group was 31 to 54 years. A total of 50% (2/4) of the participants were male, and 50% (2/4) were female. Their cardiology practice time ranged from 3 to 32 years. The session was held through videoconference using the Zoom software on June 30, 2021, and lasted 90 minutes.

For group 3, we considered eligible the 32 patients whose cases were discussed during the videoconference sessions. Unfortunately, half (16/32, 50%) of them could not be contacted due to communication hardships or other unspecified reasons. The researchers relied on the help of the community health workers from the practice for invitations. LG and MKG invited all 16 contactable patients and decided to program 1 session, forecasting a nonattendance rate of at least 30%. In total, 31% (5/16) of the invited patients and the daughter of 1 patient, who was also his caregiver, attended the meeting on July 21, 2021, at the practice's auditorium. The caregiver also contributed to the content but was identified as a patient due to privacy measures. The meeting lasted 63 minutes and was supervised by MKG, with 1 primary care physician as an observer.

For the individual interviews during phase B, we considered all 19 primary care physicians who worked as chief physicians of their respective practices in a different city region from that of the primary care practice in phase A. All accepted the invitation. A total of 79% (15/19) were women, and 84% (16/19) were White. The years of experience in primary care varied from 3 to 15 years. The interviews were conducted at the participants' workplace in the practice's lounge during work hours at a previously scheduled date and time. Importantly, medical staff and resource shortages were frequent in this region, especially during the COVID-19 pandemic, which coincided with the study period. This may have contributed to different attitudes and points of view regarding the same intervention. The interviews took place in December 2022.

The sampling for the qualitative methods was purposefully determined. The participants were considered to adequately represent the study populations as they were directly (primary care physicians and cardiologists) or indirectly (patients) involved in the telehealth process. The assessment of data saturation for the focus groups could not be planned because, despite previous consideration of repeating sessions with further participants, time constraints precluded more focus group sessions. The individual interviews had a high attendance rate (19/19, 100%), so the proposed sample was reached and considered representative of the studied population. To ensure trustworthiness, the data content from each focus group session and interview was primarily assessed as satisfactory by at least 2 researchers (MKG, LG, or PCM) at the end of each data collection activity. Due to operational reasons, transcriptions were not returned to the participants for feedback.

Data were recorded using the embedded audio recorder from LG's cellphone (iPhone SE [Apple Inc]) for the focus groups and the Telegram app (Telegram FZ-LLC) on PCM's phone for the individual interviews. All content was transcribed using the Transkriptor online platform [61] and stored locally on the investigators' PCs (LG or PCM, respectively, for the focus groups and interviews) with no online access.

### *Quantitative Methods*

In both phases of the project, we included all patients with heart failure whose cases were discussed in a telehealth session in the study. We excluded patients initially selected by the primary care physicians whose cases were not addressed in telehealth sessions. The sample size was not calculated for the quantitative assessment as hypothesis testing was not intended [56,62]. Therefore, we analyzed the baseline data of all the included participants in the study and the data after the intervention when there were enough data to be analyzed.

### *Quantitative Data*

The primary care physicians registered the clinical data from the case discussions on electronic health records. For research purposes, the teleconsultants also entered data from the telehealth sessions on a REDCap (Research Electronic Data Capture; Vanderbilt University) database [63] hosted on a secure server at the INC and accessible only to the research team. The Rio de Janeiro municipality health department granted remote access to the electronic health records to follow up on the patients.

## Data Analysis

### *Qualitative*

The transcripts were imported to the NVivo software (version 12 for transcripts from group 1 and 2 sessions and version 14 for individual interviews with physicians; Lumivero). The software version changed over the study period due to a change in license permissions by one of the research institutions [64]. MKG, LG, and PCM double-checked the content for transcription accuracy and corrected occasional mistakes in the electronically transcribed content to ensure the accuracy and confirmability of the dataset. To ensure the participants' anonymity, we identified the content by the letter corresponding to the group. We attributed *C* to cardiologists, *FP* to family physicians, *P* to patients, and *IP* to individually interviewed physicians followed by a numeral according to the order of answers within the group. We did not add notes to capture nonverbal information.

In total, 3 researchers (LG, MKG, and PCM) analyzed the transcripts using thematic analysis as the primary approach [50,65-67]. First, the authors performed a general collective reading, obtaining first impressions about the content. They then explored the content, breaking it down into sentences (units). The units were coded initially as subthemes and then classified into broader themes. The coding proceeded dynamically during the reading, driven by the content, the guiding questions, and the authors' perspectives. It was cyclical, involving rereadings until all sentences were classified. Repetitive statements were discarded. The 3 authors involved in data analysis worked together in 4 weekly in-person sessions using member checking and triangulation to enhance the analysis's credibility and dependability.

Finally, the information was summarized, enabling the critical analysis of the material from the authors' perspective. The authors emphasized the inductive interpretation of the content [65], analyzing the participants' points of view and stories rather than quantitative variables such as the frequency of themes or codes.

LG, MKG, and PCM had in-person discussions to execute the data analysis and interpretation until they reached a satisfactory consensus considering different opinions and interpretations. The contents of each focus group session and the interviews were analyzed separately.

LG, MKG, and PCM had previous professional relationships with participants in the focus groups and individual interviews. LG was the former primary care coordinator in Rio de Janeiro and had previously collaborated academically with the involved cardiologists. MKG is an associate professor at the university who runs the internship program at the primary care practice from study phase A. PCM was the medical coordinator of the group of individually interviewed primary care physicians during the study period. These factors bring critical reflexivity to the data collection and analysis as the authors are linked to the health services they study and have personal intents and

assumptions regarding assessing the study intervention, for example, the expectation of positive outcomes.

### *Quantitative*

We collected data on demography (age, sex, and race), anthropometry (weight and BMI), vital signs (blood pressure and heart rate), heart failure decompensation (defined as the presence of pulmonary rales, jugular vein stasis, or leg edema on examination), and prescribed drugs and dosage. To assess GDMT in patients with HFrEF, we considered the 3-drug regimen of renin-angiotensin-aldosterone system inhibitors (RAAS-I), β-blockers, and mineralocorticoid receptor antagonists. We observed whether the drugs were used and the target doses were reached [68]. As we collected data from 2020 to 2022, when the recommendation of sodium-glucose cotransporter 2 (SGLT-2) inhibitors in guidelines as the fourth treatment *pillar* [22,69] was not yet consolidated in medical practice or incorporated into local guidelines [68], we decided not to consider the prescription of this drug class in our assessment of GDMT. Therefore, the use of SGLT-2 inhibitors was registered but not included in the GDMT analysis.

We analyzed the data using simple descriptive statistics. We described the baseline variables of all included patients. For the subgroup of patients with follow-up data, we described and compared the proportion of patients who were decompensated. Among those, we compared the proportion of patients with HFrEF who used GDMT.

All comparisons were between baseline and the latest time point within the year after the intervention, grouped by phase. Inferential statistics were not executed because the study objective was not to test any hypothesis based on the study data. If there was more than one measurement for the same patient during follow-up, we considered only the latest time point value.

## Outcomes

The primary outcome was the feasibility of telehealth support. To draw inferences about this outcome, we integrated the qualitative exploratory findings of the content analysis of the focus groups and individual interviews with quantitative data such as patients' baseline data, clinical status, and the primary care physicians' use of GDMT. For data integration, we connected the data within selected feasibility domains described by Aschbrenner et al [70] (eg, recruitment capacity, assessment procedures, implementation resources, intervention delivery, and acceptability). For decisions about feasibility and progression to the main trial, we used the A Process for Decision-Making After Pilot and Feasibility Trials framework for feasibility analysis described by Bugge et al [71]. We presented the integration results in the form of a joint display [72].

## Ethical Considerations

This study was carried out following the Declaration of Helsinki and approved by the INC (registration 5272), the health department of the Rio de Janeiro municipality (registration 5279), the Federal University of Ouro Preto (registration 5150), and the Brazilian National Research Ethics Committee (registration 8000) under application 14894819.5.0000.5272. The assessment by the Danish Research Ethics Committee System was waived because the study did not involve Danish participants or the use of Danish data.

Patients and primary care physicians involved in the study were informed and included only after signing informed consent forms tailored to each participant category. These forms served as a formal invitation to the study explaining the rationale behind the research and detailing characteristics such as the number of participants and the study duration. We also outlined the proposed activities and disclosed the potential benefits and risks of participation. Additional topics included information on data handling and use, confidentiality, and privacy, along with clarification about involvement in the study and the absence of financial or other forms of compensation for participation.

Regarding data collection and use, the researchers sought access from the local health authority to private demographic and clinical data available in the primary care health services' electronic health record system (VitaCare). The Rio de Janeiro municipality granted authorization after we signed a statement of responsibility for data use. The informed consent permits secondary analysis without requiring additional permission.

The research team monitored patient data throughout the study. To ensure data safety, only 1 researcher and 2 undergraduate students had access to extract data from the electronic health records and input them into the study's REDCap databases. The data were pseudoanonymized, with participants identified by their national health registration numbers. The REDCap database was subsequently made available to the rest of the research team in Brazil. Case management remained unaffected except for the eventual modifications in medical decisions influenced by telehealth. All procedures adhered to relevant laws and institutional guidelines.

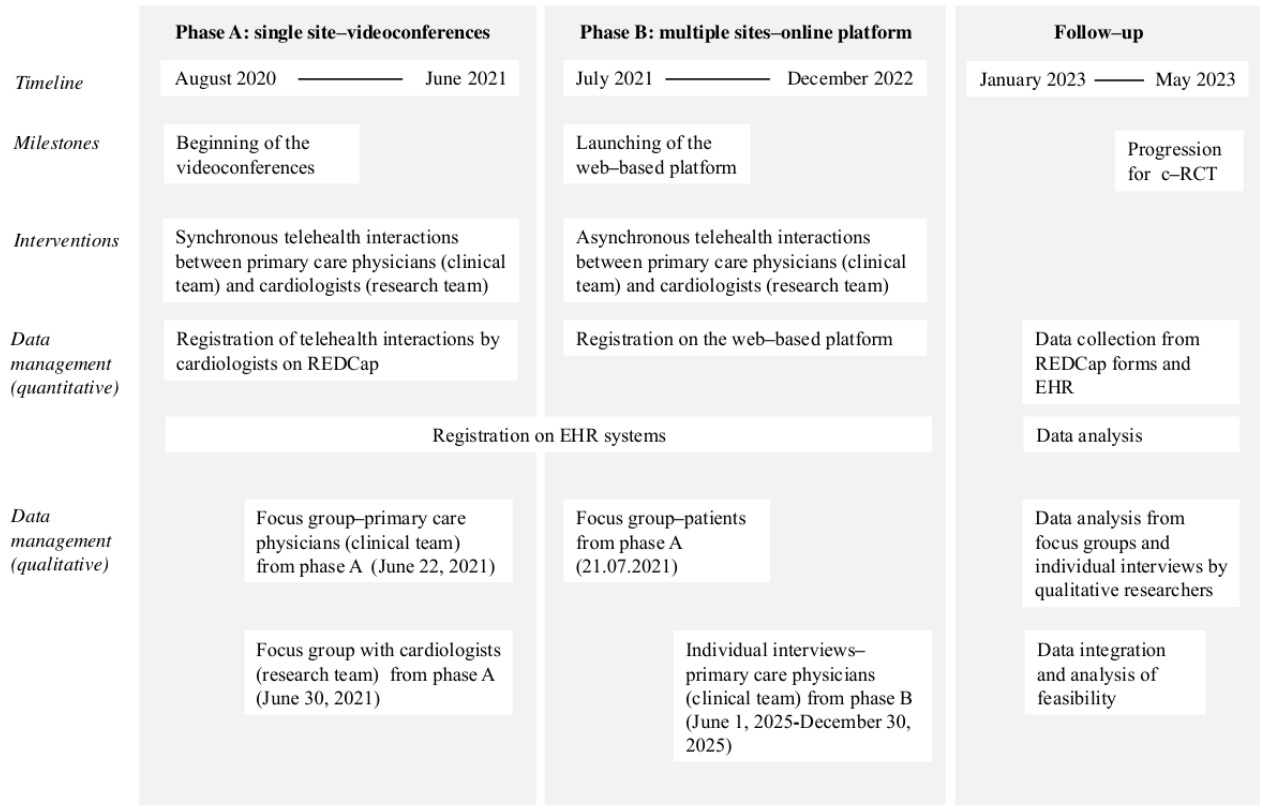
## Registration

The BRAHIT frame project is registered at ClinicalTrials.gov under the number NCT04466852 and was approved by Brazil's National Research Ethics Committee under the registration number 14894819.5.0000.5272.

## Procedural Diagram

In Figure 3, we present a procedural diagram [55] containing the timeline, the researchers' tasks, participant activities, and data collection methods according to each project phase to ensure clarity in the study methods and execution.

**Figure 3.** Procedural diagram—timeline, interventions, tasks, and data management by study phase. c-RCT: cluster-randomized controlled trial; EHR: electronic health record; REDCap: Research Electronic Data Capture.




# Results

## Qualitative Results

### *Common Findings*

The content of all qualitative activities had telehealth support as a common theme due to the specific probing questions posed to all participants. Conversely, particular themes emerged based on the participant categories. For instance, concerns about patients' social conditions and interactions among health care sectors were highlighted among primary care physicians in phase A but were less evident among those in phase B, where the themes focused more on professional matters. Differences in physicians' educational backgrounds may explain this variation. All primary care physicians in phase A (focus group; 15/15, 100%) specialized in family and community medicine, whereas only 37% (7/19) in phase B (individual interviews) had the same specialization.

On the other hand, the physicians interviewed in the project's phase B were more experienced than the ones in phase A. Different data collection methods (interview vs focus group) could have also played a role. In the case of the cardiologists, the operational aspects were notably frequent, which correlates with the fact that they were the consultants and research team members. In the patient focus group, the themes actively mentioned by the participants were related to the primary care service organization and their experience with disease and care. Each group's code classification, findings, and interpretation are detailed in the following sections.

### *Focus Group: Primary Care Physicians*

#### Overview

Four themes emerged from the session's content analysis: (1) population aspects, (2) clinical competence in primary care, (3) communication among health care services, and (4) telehealth support. The themes, subthemes, and definitions are shown in Table 1.

**Table 1.** Focus group 1 (primary care physicians)—themes, subthemes, and definitions that emerged from content analysis.

| Theme and subtheme | Definition |
|---|---|
| **Population aspects** | |
| Disparities | Opinion on the population's socioeconomic and cultural vulnerability |
| Mobility | Patients' mobility hardships |
| **Clinical competence in primary care** | |
| Confidence | Lack of confidence in managing patients with heart failure |
| Task perception | Perception of the task of treating patients with heart failure |
| Communication among health care services | Communication gap among health care sectors |
| **Telehealth support** | |
| Use | Discussion about the use of supporting tools |
| Potential and barriers | Assessment of telehealth support use |

### Population Aspects

Considering the context in which the focus group took place, a socially deprived area of the city, and the educational background of the participants, who were trained to deliver person-centered, community-oriented care, the mention of social disparities and their impact on patient care and the service organization was expected. The discussion highlighted the population's socioeconomic and cultural vulnerability, which markedly influences their lives and clinical follow-up [73]:

> *...our patients are very vulnerable...So economically, intellectually, and culturally speaking, they need us.* [FP1]

Another important subtopic was *mobility*, reflecting the concerns of the primary care physicians about the patients' itinerary within and between health care services. The patients' difficulties moving around the city for an eventual referral to a specialized service were reported, reinforcing the importance of the primary care practice offering close, accessible, and comprehensive care, facilitating adherence. This aspect is supported by findings from the literature correlating the accessibility of primary care facilities and its impact on the continuity and quality of primary care delivery [74,75]:

> *...They don't have the financial conditions to do it (commuting) from their pocket. So, they will return to us to continue care.* [FP1]

### Clinical Competence in Primary Care

An essential theme that emerged from this focus group was the primary care physicians' confidence in assisting patients with cardiologic conditions such as heart failure. The lack of confidence reported by some physicians regarding themselves and their colleagues may be due to inexperience and insufficient training before graduation:

> *...We know some topics more basically, like reading an X-ray or an electrocardiogram. I think the EKG is a general difficulty.* [FP3]

There was also sometimes a notably unclear perception of primary care as a scenario for managing severe diseases such as heart failure:

> *...I always imagined that I would manage...here in primary care, only hypertension, so anything that goes a little beyond within cardiology topics, literally, I don't know.* [FP2]

### Communication Among Health Care Services

When collaborative care is discussed, one main topic that usually emerges is the communication hardships between services [76]. The participants described significant communication problems, which led to gaps and unawareness of actions performed in secondary and tertiary services, affecting the patients' care:

> *...I think the great difficulty we have today is that we seldom receive a report from a specialist. They should tell us how shared care is supposed to happen...* [FP1]

> *Sometimes, they order tests or prescribe medication, and we don't know exactly why. How can I share the care with them and continue if I don't know where they want to go?* [FP3]

### Telehealth Support

The researchers' questions probed the ubiquitous theme of teleconsulting services. The group discussed the ideal characteristics of a teleconsulting service, their experience with the BRAHIT project, and other support activities. The group evaluated telehealth support positively as it was easily accessible. They also assessed the BRAHIT project as having favorable characteristics:

> *...the intimacy, the ability (of the teleconsultants) to understand my difficulty, because sometimes I ask a question, and he already answers...* [FP9]

> *...They are focal specialists who understand my reality and see that they are contributing not only to me, but to patient care.* [FP5]

On the other hand, the time-consuming effort required to be physically present during the videoconferences was a frequent negative feedback. This information led the researchers to refine the intervention, adapting the telehealth offer to include an asynchronous approach commonly used in other telehealth services [77]:

*...We know that we are privileged, because there are a lot of physicians here, but in other clinics I have worked, I would rarely have the time to be online in a web conference.* [FP10]

### *Focus Group: Cardiologists*

#### Overview

Two themes emerged from the session's content analysis: (1) the relationship with the primary care service and (2) telehealth support. The themes, subthemes, and definitions are shown in Table 2.

**Table 2.** Focus group 2 (cardiologists)—themes, subthemes, and definitions that emerged from content analysis.

| Theme and subtheme | Definition |
|---|---|
| **Relationship with the primary care service** | |
| Vision on primary care | Discussion about their vision on primary care services |
| Mission | The National Institute of Cardiology's mission as a teaching institution |
| **Telehealth support** | |
| Education | Evaluation of the interactions regarding collaboration |
| Challenges | Challenges of telehealth implementation |

#### Relationship With the Primary Care Service

The cardiologists discussed their preconception about primary care services, initially evaluated as deficient in structure and quality of human resources, and stated a paradigm shift after contact with the team from the primary care practice:

*...we are hospitalists, and sometimes we believe that the primary care practice has an inadequate structure, right?* [C1]

*Sometimes, physicians do not have adequate training, and it was a paradigm that was broken about the technical level of the colleagues, which is, in fact, very high.* [C2]

Another important finding was the recognition by the cardiologists of significant opportunities for the INC team, highlighting their role as a specialized public institution in education to improve the overall quality of the health care system:

*...I noticed since the first time the chance not only to improve the follow-up of these patients but also to teach the professionals who work there, allowing them to feel more capable of helping people. I think that most people in primary care have this vocation.* [C1]

#### Telehealth Support

The telehealth interactions were assessed as positive regarding training and collaboration between the parties, and opportunities for bilateral learning were identified:

*They already have a different perception of approaching cardiac patients, and it has been a very enriching exchange of experiences for both sides. Sometimes, I think we also learn from them.* [C2]

*So, bringing not only knowledge but also the experience that we have in terms of treatment, I think general practitioners have good experiences with us and realize that we are calm. The patient is severe, but we manage it.* [C4]

The cardiologists reported concerns about implementing telehealth, specifically about its scalability and sustainability and the engagement of primary care physicians:

*...I just think there was also an underuse of the service. I think it could have been used more.* [C2]

### *Focus Group: Patients*

#### Overview

Two themes emerged from the session's content analysis: (1) disease and care experience and (2) telehealth support. The themes, subthemes, and definitions are shown in Table 3.

**Table 3.** Focus group 3 (patients—phase A)—themes, subthemes, and definitions that emerged from content analysis.

| Theme and subtheme | Definition |
|---|---|
| **Disease and care experience** | |
| Health literacy | Understanding regarding their disease and care |
| Insights about self-care | Thoughts about good habits and well-being |
| Care evaluation | Assessment of physicians' actions and consequences for their health |
| Free will | Attitudes toward the disease |
| **Telehealth support** | |
| Opinions and fears | Opinions and worries about telehealth support |

### Disease and Care Experience

The probing questions for the patients investigated their understanding of heart failure as a disease and their conceptions of medical assistance. Their discussions revealed a heterogeneous understanding of cardiologic conditions and their treatment:

> *...I used to think there was one type of heart disease. One would feel chest pain. But it seems that there is more than that. I do not understand.* [P3]

There were also reports about the patients' improvements after they were properly diagnosed and treated. They could find a positive correlation between following correct habits and taking correct medications and their well-being:

> *...Then I do not feel tired anymore. It has been two years now. I cycle to work and to everywhere around. I help a friend with construction work. It is impressive. I even get suspicious sometimes.* [P6]

Nevertheless, in the words of other participants, we recognized a disconnection between their interpretation of physicians' actions, test results, and medications and their feelings. We also noticed different attitudes toward the disease depending on individual characteristics:

> *...I only go to hospitals or clinics if I am dying. If I feel something that can be managed with analgesics or something, I will not come. I do not take prescription medications every day, as I feel myself controlled.* [P4]

### Telehealth Support

The participants responded positively when discussing cardiologists' telehealth support for their primary care physicians. They understood the initiative as an improvement. One participant reported that his physician participated in the BRAHIT project:

> *...He [the physician] takes pictures of the test results and sends them to the project. Yes, I think he is participating. Maybe it is working!* [P6]
>
> *...I think it is a very good idea.* [P2]

The literature does not extensively address the patient vision of telehealth between health care professionals. Our findings are significant as they provide the patients' perspective on the strategy. In our findings, the patients seen in specialized care reported feeling unsafe enough to stop regularly attending specialist appointments even after the implementation of telehealth support:

> *...I think it would be better if we went to the hospital and had all the tests. It would be better to go directly there. Because it is a specialist.* [P3]
>
> *...I go to the hospital every three months. I feel safer going there, too.* [P5]

## Individual Interviews

### Overview

Four themes emerged from the interview content analysis: (1) work overload, (2) telehealth use, (3) clinical competence, and (4) referral practices. The themes, subthemes, and definitions are shown in Table 4.

**Table 4.** Individual interviews (primary care physicians—phase B)—themes, subthemes, and definitions that emerged from content analysis.

| Theme and subtheme | Definition |
|---|---|
| Work overload | Influence of work rhythm on telehealth use |
| **Telehealth support** | |
| Actual use | Experiences using telehealth |
| Barriers | Reasons for not using telehealth |
| Clinical competence | Confidence in assisting patients with heart failure |
| Referral practices | Influence of telehealth in referring patients to specialists |

### Work Overload

Professionals usually describe the work context in Brazil's primary care practices as being in high demand. Most practices have a high panel size, and the teams usually must deal with acute and programmed care. The scenario during our research was influenced by the COVID-19 pandemic, bringing further pressure to the practices and the political scene, where the Rio de Janeiro municipality was adopting an austerity policy, including staff reduction, which also played a role [78-80]. Therefore, the principal issue reported by the participants was the lack of available time due to an overwhelming burden of tasks and consultations:

> *We did not use the telehealth support because of the work overload in our practice, a significant physician shortage, and turnover. This jeopardized the dissemination and utilization of the tool.* [IP2]

### Telehealth Support

Some participants reported a favorable experience and advantages, such as greater confidence in managing patients with heart failure and fewer referrals. They recognized the initiative's potential for quality improvement:

> *...discussing cases of patients with heart failure with multimorbidity and decompensated cases provided greater confidence in managing the case and could reduce referrals to emergencies and specialists.* [IP1]

Conversely, cardiologists sometimes took a long time to respond to contact requests, which was considered a problem:

> *When I tried to use the website, connecting was hard. I found it slow. As other tools are available online, I do not use them anymore.* [IP3]

### Clinical Competence

When asked about their ability and confidence in assisting patients with heart failure, most physicians answered that they could help. This finding brings about an interesting paradox because our quantitative data showed a poor clinical baseline status of most patients whose cases were discussed in the project:

> *...no need for questioning in cardiology; therefore, I have not used the telehealth support from the BRAHIT project. It is worth mentioning that we have a WhatsApp group for case discussions provided by the municipality health department.* [IP5]

Other reports mentioned a lack of interest, use of alternative tools, or no need to use telehealth support:

> *...in my population, there are no patients with heart failure needing specialist consultation, nor do I need telehealth support for myself.* [IP6]

### Referral Practices

The traditional approach to treating complex cases in primary care involves referring patients to specialized services. A total of 16% (3/19) of the participants alleged that referring the patient to the cardiology service would be easier. Nevertheless, this approach may entail problems, such as low patient attendance due to the issues described previously, such as commuting difficulties, which are also reported in the literature [5,81,82]:

> *...When I need to refer the patient to a cardiologist, I use the referral system. So, the telehealth support offer and objectives are still not clear to me.* [IP7]

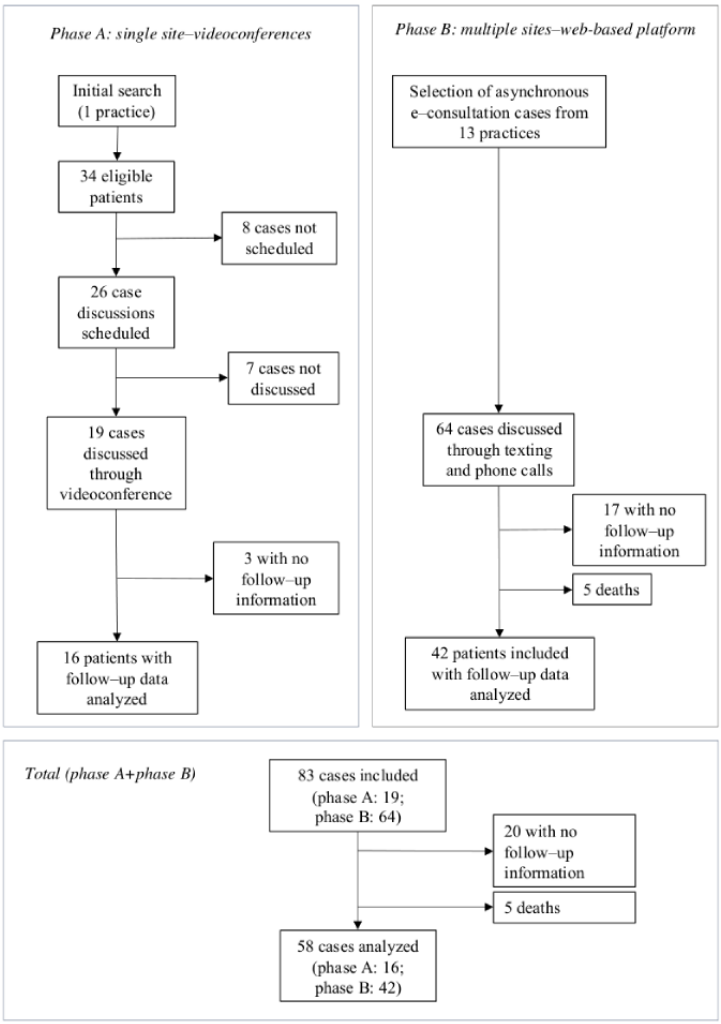
> *...The patients have already been managed via referral through the referral system.* [IP8]

## Quantitative Results

### *Participants*

During the videoconference phase (phase A) of the intervention, the physicians selected 34 patient cases for discussion, of which 26 (76%) were scheduled for discussion based on the physicians' criteria and their availability to attend the telehealth session. A total of 27% (7/26) of these cases were not discussed for unknown reasons. In total, 73% (19/26) of the cases were discussed via videoconference. Follow-up data were available from the practice's electronic health records for 84% (16/19) of these patients. In phase B, 64 patients from 13 primary care practices had their cases discussed asynchronously. Of these 64 patients, 5 (8%) died, 17 (27%) did not have further consultation records, and the remaining 42 (66%) were followed up on. Adding both phases, 83 cases were discussed, and 58 (70%) patients were followed up on. Participant inclusion is summarized in the flowchart in Figure 4.

**Figure 4.** Flow diagram of patient inclusion in the study and quantitative before-and-after follow-up for 1 year based on the CONSORT (Consolidated Standards of Reporting Trials) framework for reporting clinical trials (data from August 2020 to December 2022).




### *Baseline Data*

Regarding demographic data, the mean patient age was 61 (SD 12) years. Of the 83 patients, 52 (63%) were male, and 31 (37%) were female; of 73 patients with available data, 30 (41%) were White, and 28 (38%) were Black or belonged to another ethnic minority group. The proportion of common diagnoses associated with heart failure was similar to that in the literature except for chronic obstructive pulmonary disease, which was reported in only 2% (1/61) of the participants with available data, suggesting underdiagnosis [83]. Regarding anthropometry and vital signs, BMI and mean blood pressure and heart rate values were above the recommended limits. Of the patients with available data, 64% (7/11) in phase A and 45% (21/47) in phase B had HFrEF. Most patients (39/74, 53%) had poor physical status according to the New York Heart Association classification. The data are described in detail in Table 5.

**Table 5.** Baseline demographic and clinical data of all patients included in the quantitative assessment of this study (N=83).

| Variable | Phase A (n=19) | Phase B (n=64) | Total |
|---|---|---|---|
| Age (y), mean (SD; range) | 58 (12; 35-76) | 61 (13; 37-89) | 61 (13; 35-89) |
| **Sex, n (%)** | | | |
| Female | 7 (37) | 24 (37) | 31 (37) |
| Male | 12 (63) | 40 (63) | 52 (63) |
| **Race, n (%)** | | | |
| Black or other ethnic minority group | 6 (35) | 22 (39) | 28 (38) |
| White | 9 (53) | 21 (38) | 30 (41) |
| Not informed | 2 (12) | 13 (23) | 15 (21) |
| Missing | 2 (11) | 8 (12) | 10 (12) |
| **Atrial fibrillation, n (%)** | | | |
| No | 13 (76) | 33 (70) | 46 (72) |
| Yes | 4 (24) | 14 (30) | 18 (28) |
| Missing | 2 (11) | 17 (27) | 19 (23) |
| **Diabetes, n (%)** | | | |
| No | 11 (65) | 32 (62) | 43 (62) |
| Yes | 6 (35) | 20 (38) | 26 (38) |
| Missing | 2 (11) | 12 (19) | 14 (17) |
| **COPD[a] , n (%)** | | | |
| No | 13 (93) | 47 (100) | 60 (98) |
| Yes | 1 (7) | 0 (0) | 1 (2) |
| Missing | 5 (26) | 17 (27) | 22 (27) |
| **Coronary artery disease, n (%)** | | | |
| No | 7 (88) | 22 (49) | 29 (55) |
| Yes | 1 (12) | 23 (51) | 24 (45) |
| Missing | 11 (58) | 19 (30) | 30 (36) |
| **Hypertension, n (%)** | | | |
| No | 4 (21) | 15 (25) | 19 (24) |
| Yes | 15 (79) | 46 (75) | 61 (76) |
| Missing | 0 (0) | 3 (5) | 3 (4) |
| **Stroke, n (%)** | | | |
| No | 15 (100) | 56 (92) | 71 (93) |
| Yes | 0 (0) | 5 (8) | 5 (7) |
| Missing | 4 (21) | 3 (5) | 7 (8) |
| **Peripheral artery disease, n (%)** | | | |
| No | 15 (100) | 56 (97) | 71 (97) |
| Yes | 0 (0) | 2 (3) | 2 (3) |
| Missing | 4 (21) | 6 (9) | 10 (12) |
| **Dyslipidemia, n (%)** | | | |
| No | 8 (62) | 26 (53) | 34 (55) |
| Yes | 5 (38) | 23 (47) | 28 (45) |
| Missing | 6 (32) | 15 (23) | 21 (25) |

| Variable | Phase A (n=19) | Phase B (n=64) | Total |
|---|---|---|---|
| BMI (kg/m$^2$), mean (SD; range) | 32 (7; 23-49)[b] | 29 (6; 19-53)[b] | 30 (6; 19-53)[c] |
| Systolic blood pressure (mm Hg), mean (SD; range) | 138 (31; 97-220) | 130 (29; 90-240)[b] | 132 (29; 90-240)[b] |
| Diastolic blood pressure (mm Hg), mean (SD; range) | 91 (21; 60-160) | 80 (16; 40-120)[b] | 82 (18; 40-160)[b] |
| Heart rate (bpm[d]), mean (SD; range) | 81 (19; 53-125)[b] | 79 (18; 42-121)[b] | 79 (18; 42-125)[c] |
| **NYHA[e] functional classification, n (%)** | | | |
| I | 2 (12) | 10 (17) | 12 (16) |
| II | 8 (50) | 15 (26) | 23 (31) |
| III | 1 (6) | 21 (36) | 22 (30) |
| IV | 5 (31) | 12 (21) | 17 (23) |
| Missing | 3 (16) | 6 (9) | 9 (11) |
| LVEF[f] (%), mean (SD; range) | 35 (8; 21-48)[g] | 43 (19; 14-80)[h] | 42 (18; 14-80)[i] |
| **Heart failure classification (LVEF status), n (%)** | | | |
| Reduced | 7 (64) | 21 (45) | 28 (48) |
| Mildly reduced | 4 (36) | 9 (19) | 13 (22) |
| Preserved | 0 (0) | 17 (36) | 17 (29) |
| Missing | 8 (42) | 17 (27) | 25 (30) |
| Creatinine (mg/dL), mean (SD; range) | 1.3 (1; 0.7-5.1)[c] | 1.3 (1; 0.6-8.0)[j] | 1.3 (1; 0.6-8.0)[k] |

[a]COPD: chronic obstructive pulmonary disease.
[b]Missing: n=1.
[c]Missing: n=2.
[d]bpm: beats per minute.
[e]NYHA: New York Heart Association.
[f]LVEF: left ventricular ejection fraction.
[g]Missing: n=8.
[h]Missing: n=17.
[i]Missing: n=25.
[j]Missing: n=3.
[k]Missing: n=5.

### *Outcome Analysis*

We used data from 58 patients available in electronic health records within 1 year following the first telehealth interaction to assess changes before and after telehealth. The mean follow-up time after telehealth was 183 (SD 109; range 14-365) days. The proportion of missing data at follow-up was very high (mean 28%, SD 14%, varying from 1/21, 5% to 23/42, 55% depending on the variable), precluding a precise assessment or identification of patterns.

There was a modest change in the patients' vital signs after follow-up compared to baseline. The mean systolic blood pressure was 7 mm Hg lower, the mean diastolic blood pressure was 3 mm Hg lower, and the mean heart rate was 3 beats per minute lower. The proportion of patients with signs of decompensated heart failure was 63% (17/27) compared to 50% (29/58) of patients at baseline. Of the patients with reduced ejection fraction assessed at baseline and during follow-up, 55% (12/22) and 55% (11/20), respectively, had prescriptions for the 3 main GDMT drug classes, which can be explained by an increase in β-blocker (17/20, 85% to 18/19, 95%) and RAAS-I (14/20, 70% to 15/19, 79%) prescription but a drop in the prescription of spironolactone (16/20, 80% to 15/20, 75%). Newer agents such as neprilysin and SGLT-2 inhibitors were introduced during the follow-up period for 4 and 1 patient, respectively, compared to no use record at baseline. The data are presented in detail in Table 6.

**Table 6.** Clinical data before and after telehealth support—subgroup of patients with at least one follow-up contact registered in primary care electronic health records (N=58).

| Variable | Phase A (n=16) | | Phase B (n=42) | | Total | |
|---|---|---|---|---|---|---|
| | Before | After | Before | After | Before | After |
| Days between baseline and follow-up, mean (SD; range) | 157 (109; 14-344) | __[a] | 192 (99; 22-365) | — | 183 (103; 14-365) | — |
| **Heart failure classification (LVEF[b] status), n/N (%)** | | | | | | |
| Reduced | 7/9 (78) | 7/9 (78) | 15/31 (48) | 15/31 (48) | 22/40 (55) | 22/40 (55) |
| Mildly reduced | 2/9 (22) | 2/9 (22) | 4/31 (13) | 4/31 (13) | 6/40 (15) | 6/40 (15) |
| Preserved | 0/9 (0) | 0/9 (0) | 12/31 (39) | 12/31 (39) | 12/40 (30) | 12/40 (30) |
| Missing | 7/16 (44) | 7/16 (44) | 11/42 (26) | 11/42 (26) | 18/58 (31) | 18/58 (31) |
| Systolic blood pressure (mm Hg), mean (SD; range) | 136 (33; 97-220) | 134 (43; 90-260)[c] | 132 (31; 90-240) | 123 (22; 70-160)[d] | 133 (32; 90-240) | 126 (30; 70-260)[e] |
| Diastolic blood pressure (mm Hg), mean (SD; range) | 91 (23; 60-160) | 88 (21; 60-140)[c] | 80 (17; 40-120) | 77 (16; 40-109)[d] | 83 (19; 40-160) | 80 (18; 40-140)[e] |
| Heart rate (bpm[f]), mean (SD; range) | 85 (19; 58-125)[g] | 86 (20; 63-125)[h] | 79 (18; 42-120) | 74 (13; 43-100)[i] | 80 (18; 42-125)[g] | 77 (15; 43-125)[j] |
| **Signs of decompensated heart failures[k] , n/N (%)** | | | | | | |
| No | 5/14 (36) | 5/8 (62) | 18/38 (47) | 5/19 (26) | 23/52 (44) | 10/27 (37) |
| Yes | 9/14 (64) | 3/8 (38) | 20/38 (53) | 14/19 (74) | 29/52 (56) | 17/27 (63) |
| Missing | 2/16 (12) | 8/16 (50) | 4/42 (10) | 23/42 (55) | 6/58 (10) | 31/58 (53) |
| **GDMT[l] in HFrEF[m,n] , n/N (%)** | | | | | | |
| No | 4/7 (57) | 3/7 (43) | 6/15 (40) | 6/13 (46) | 10/22 (45) | 9/20 (45) |
| Yes | 3/7 (43) | 4/7 (57) | 9/15 (60) | 7/13 (54) | 12/22 (55) | 11/20 (55) |
| Missing | 0/7 (0) | 0/7 (0) | 0/15 (0) | 2/15 (13) | 0/22 (0) | 2/22 (9) |
| **β-blocker use in HFrEF, n/N (%)** | | | | | | |
| No | 2/7 (29) | 0/7 (0) | 1/13 (8) | 1/12 (8) | 3/20 (15) | 1/19 (5) |
| Yes | 5/7 (71) | 7/7 (100) | 12/13 (92) | 11/12 (92) | 17/20 (85) | 18/19 (95) |
| Missing | 0/7 (0) | 4/11 (36) | 0/13 (0) | 4/16 (25) | 0/20 (0) | 8/27 (30) |
| **MRA[o] use in HFrEF, n/N (%)** | | | | | | |
| No | 3/7 (43) | 3/8 (38) | 1/13 (8) | 2/12 (17) | 4/20 (20) | 5/20 (25) |
| Yes | 4/7 (57) | 5/8 (62 | 12/13 (92) | 10/12 (83) | 16/20 (80) | 15/20 (75) |
| Missing | 1/8 (12) | 4/12 (33) | 0/13 (0) | 3/15 (20) | 1/21 (5) | 7/27 (26) |
| **RAAS-I[p] use in HFrEF, n/N (%)** | | | | | | |
| No | 3/7 (43) | 1/7 (14) | 3/13 (23) | 3/12 (25) | 6/20 (30) | 4/19 (21) |
| Yes | 4/7 (57) | 6/7 (86) | 10/13 (77 | 9/12 (75) | 14/20 (70) | 15/19 (79) |
| Missing | 0/7 (0) | 4/11 (36) | 0/13 (0) | 2/14 (14) | 0/20 (0) | 6/25 (24) |
| **Neprilysin inhibitor use in HFrEF, n/N (%)** | | | | | | |
| No | 7/7 (100) | 5/8 (62) | 13/13 (100) | 11/12 (92) | 20/20 (100) | 16/20 (80) |
| Yes | 0/7 (0) | 3/8 (38) | 0/13 (0) | 1/12 (8) | 0/20 (0) | 4/20 (20) |
| Missing | 2/9 (22) | 5/13 (38) | 0/13 (0) | 3/15 (20) | 2/22 (9) | 8/28 (29) |
| **SGLT-2[q] inhibitor use in HFrEF, n (%)** | | | | | | |
| No | 7/7 (100) | 7/8 (88) | 12/12 (100) | 12/12 (100) | 19/19 (100) | 19/20 (95) |
| Yes | 0/7 (0) | 1/8 (12) | 0/12 (0) | 0/12 (0) | 0/19 (0) | 1/20 (5) |

| Variable | Phase A (n=16) | | Phase B (n=42) | | Total | |
|---|---|---|---|---|---|---|
| | Before | After | Before | After | Before | After |
| Missing | 2/9 (22) | 5/13 (38) | 1/13 (8) | 3/15 (20) | 3/22 (14) | 8/28 (29) |

[a]Not applicable.

[b]LVEF: left ventricular ejection fraction.

[c]Missing: n=5.

[d]Missing: n=13.

[e]Missing: n=18.

[f]bpm: beats per minute.

[g]Missing: n=1.

[h]Missing: n=7.

[i]Missing: n=19.

[j]Missing: n=26.

[k]Pulmonary rales, jugular stasis, or leg edema.

[l]GDMT: guideline-directed medical therapy.

[m]HFrEF: heart failure with reduced ejection fraction.

[n]GDMT—at least one renin-angiotensin-aldosterone system inhibitor+1 β-blocker+1 mineralocorticoid antagonist.

[o]MRA: mineralocorticoid receptor antagonist.

[p]RAAS-I: renin-angiotensin-aldosterone system inhibitor.

[q]SGLT-2: sodium-glucose cotransporter 2.

## Data Integration and Feasibility Assessment

The content analysis of the focus groups and individual interviews gave us a clear view of the intervention context, allowing us to identify some patterns. While assessing the feasibility of the intervention, we received critical feedback. We obtained significant insights on the implementation context and potential barriers and facilitators for the planned intervention to be appropriately delivered within the upcoming cluster-randomized trial. In turn, the quantitative analysis showed the baseline status regarding the patients' demographics and clinical characteristics and some change tendencies in the primary care physicians' prescription practices after telehealth implementation.

To draw inferences about both data types, we interconnected the main findings and correlated them with feasibility domains [70] when applicable. We concluded that the intervention is feasible, with adjustments, as described in the A Process for Decision-Making After Pilot and Feasibility Trials model items *adapting the intervention*, *adjusting the clinical context within which the intervention would be delivered*, and *amending elements of the trial design* [71]. Practically, during the feasibility trial, we decided to use the asynchronous telehealth method and recruit patients discharged from hospitals and emergency rooms in the future cluster-randomized trial instead of only including the patients selected by the primary care physicians. Table 7 consolidates the main findings, interpretations, and decisions regarding feasibility in a joint display.

**Table 7.** Joint display of results and mixed methods interpretations integrating qualitative and quantitative findings.

| Domain | Quantitative results | Qualitative results | Mixed methods interpretation | ADePT[a] actions |
|---|---|---|---|---|
| Setting | • Of 73 patients with available data, 30 (41%) were White, and 28 (38%) were Black or from other ethnic minority groups, contrasting with the population of the study.<br>• The mean age of the study participants was 61 years, 4.5 years lower than the mean reported age in Brazil of patients with heart failure. | • Primary care teams reported lack of physicians in individual interviews.<br>• The population covered by the practice is socioeconomically vulnerable and has insufficient knowledge about their condition and care. | • The setting is challenging, requiring active involvement of all stakeholders.<br>• Facing difficulties, physicians may privilege patients with easier access to care.<br>• Actions integrated with telehealth support aimed at patient health literacy could be synergic. | • Adapt the intervention for the setting conditions.<br>• Be aware of possible access hardships for non-White populations.<br>• Design cointerventions to overcome barriers (eg, patient education activities). |
| Recruitment capacity | • A total of 83 patients had their cases discussed in 2 years in the practices where physicians used the telehealth offer.<br>• Only 1 in 15 physicians who participated in the individual interviews used the telehealth offer. | • Lack of awareness on the part of the primary care physicians of their need for support.<br>• Work overload hindered the use of cardiologist support with telehealth. | • The results agree and are likely to have a strong correlation.<br>• An active search by the research team of patients suitable for telehealth could help. | • Modifying the intervention to include a nudging strategy for telehealth use would favor recruitment.<br>• A decision was made to include actively sought out postdischarge patients in the subsequent trial. |
| Assessment procedures | • Identification of improvement opportunities from the baseline clinical data<br>• Use rate of newer agents to treat heart failure improved from 0 (0%) to 5 (20%).<br>• Lack of effect in other quantitative outcomes (eg, patients who were decompensated) | • Both teleconsultant cardiologists and family physicians are optimistic about using telehealth as a tool for care improvement.<br>• Lack of awareness of support need by some primary care physicians related to the telehealth offer | • The results agree and are likely to have a strong correlation. | • The intervention is feasible and potentially beneficial for the clinical performance.<br>• Design cointerventions to overcome barriers (eg, professional education activities). |
| Intervention delivery | • Identification of improvements related to the intervention<br>• Use rate of newer agents to treat heart failure improved from 0 (0%) to 5 (20%). | • Positive feedback from the participants from the primary care teams<br>• Videoconferences were time-consuming. | • The results agree and are likely to have a correlation. | • The intervention is feasible if adapted. The intervention was modified for asynchronous communication in phase B. |
| Implementation resources | • The upscaled offer of telehealth was rapidly accepted in 13 primary care practices in phase B.<br>• The telehealth offer seemed cost-effective and did not cause a burden to the project finances. | • The feedback from teleconsultants was positive.<br>• The sustainability of the offer was a concern in the cardiologist focus group. | • The results agree and are likely to have a correlation. | • The intervention is feasible. |

| Domain | Quantitative results | Qualitative results | Mixed methods interpretation | ADePT[a] actions |
|---|---|---|---|---|
| Acceptability | • There was no refusal from primary care physicians to participate in the study, although compliance with the intervention was low in some settings. | • Content analysis of the patient focus group revealed restrictions regarding the intervention as it could be a risk for prompt access to specialized care. | • There was an attention point regarding the guarantee of access to specialized care. | • The intervention can be tailored to include clarification about no access block for the patients. |

[a]ADePT: A Process for Decision-Making After Pilot and Feasibility Trials.

## Discussion

### Principal Findings and Interpretation

In this study, we aimed to assess the feasibility of telehealth support from cardiologists to primary care physicians for the care of patients with heart failure in the community setting. We analyzed factors from the study context, stakeholders' attitudes and perceptions, barriers, facilitators, and possible influence on clinical practice.

The content analysis from focus groups and individual interviews revealed a favorable opinion when participants were asked about telehealth. In parallel, aspects of the intervention's context emerged, such as the population's socioeconomic conditions and primary care professionals' work environment, collaboration with other health care sectors, and professional educational background. Considering these aspects and others that may ensue in different contexts is vital while implementing and assessing telehealth interventions, as in any innovation strategy.

The assessment of context and human factors has been described as essential in several publications about social, complexity, and implementation science. Therefore, the findings of this feasibility study are consistent with the literature on complex interventions involving knowledge-seeking behavior, including eHealth technologies. In a review about spreading and scaling innovation and improvement, Greenhalgh and Papoutsi [42] add *develop adaptive capability in staff*, *attend to human relationships*, and *harness conflict productively* as principles to be followed when planning the change programs described by Lanham et al [84]. Other reviews and editorials by Robert et al [41], Greenhalgh et al [42,43], and Greenhalgh and Russell [85] refer to some hardships that we also found in our study.

Phase B participants who were interviewed reported low engagement and acceptance due to work overload. The findings echo some reports in the literature. One specific scoping review on shared decision-making strategies using digital health technology in cardiovascular care points to *increased work responsibilities* as the most frequently reported barrier [86]. The low perception of the relative advantage of telehealth, present in the analysis of individual interviews, can hinder the implementation of innovations and, therefore, must be addressed and discussed before the implementation of telehealth [87]. This finding contrasts with recent surveys about continuing medical education in primary care, where the most frequent reasons for low engagement, in addition to work overload, were the inability to use digital tools and the difficulty in integrating the process into the practice routine [88].

Another key finding was the patients' preoccupation that telehealth support could block their access to specialized services. This points to the need to reassure the patients that access to the focal specialists will still be available when using telehealth. The literature does not usually describe the patients' perspective on provider-to-provider telehealth. We believe that including their assessment is essential and highly recommended in feasibility studies [89].

Regarding demographic data, the patients' mean age was 4.5 years lower than the Brazilian average reported by the National Brazilian Registry of Heart Failure [90]. We believe that the participants' low socioeconomic status plays a role in this disparity. Studies show an earlier and higher exposure to suboptimal nutrition habits and low self-care in socially deprived populations, anticipating the development of risk factors and diseases that will cause heart failure [73,91]. There was also a low proportion of participants who were female, Black, and of other ethnic minority groups in this study, contrasting with the more frequent use of health care services by women [92] and the higher heart failure prevalence among Black people and those of other ethnic minority groups [93]. The demographic profile of our sample may indicate a selection bias by the primary care physicians when including the patients for case discussion. This finding is supported by other authors describing equity discrepancies and underrepresentation of minority groups regarding access to care [94] and research participation [95].

The quantitative analysis showed opportunities for improvement in patient care. At baseline, more than half (39/74, 53%) of the patients with available data had poor functional capacity. The low rate of GDMT use may be a reason as only 55% (12/22) of the patients with HFrEF had prescriptions according to the recommended local and international guidelines. Unfortunately, this phenomenon is frequently reported in the medical literature [8,25,69,96]. We evaluate the tendency toward GDMT as favorable, with increases in the use of all drug classes except spironolactone, whose prescription decreased. Possible reasons include variations in drug availability in primary care, as physicians usually prescribe what is available for the patients to collect for free in the practices, or the primary care physicians' lack of familiarity with the drug. The Change the Management of Patients With Heart Failure registry published by Greene et al [24] showed that mineralocorticoids were the least prescribed drug among the 3 categories (not prescribed in 67% of the patients vs 27% and 33% of the patients not being

prescribed RAAS-I and β-blockers, respectively). However, the small number of participants assessed for this outcome does not allow us to draw accurate conclusions.

Integrating qualitative and quantitative data allowed us to foresee elements to be tailored in the forthcoming clinical trial as we evaluated its context, stakeholders' attitudes, and other practicalities. We deemed the feasibility analysis positive considering the adjustments and complementary strategies within the research's reach. Accordingly, we changed the recruitment strategy, selecting patients discharged from hospitals and emergency rooms because of heart failure instead of depending on primary care physicians' spontaneous use of telehealth. We also defined the asynchronous telehealth model as the intervention and planned the implementation of educational activities to engage the target stakeholders [46].

## Strengths

This study's strength lies in its use of mixed methods to analyze data integration between the participants' opinions and the possible changes caused by telehealth. Mixed methods are recommended for studying the feasibility of complex interventions such as telehealth [48]. Integrating qualitative and quantitative data allows for a more thorough description of the intervention's development and provides specific answers for researchers, allowing for a better assessment of the feasibility domains [57,70]. Another strength was using a particular framework for decision-making in feasibility trials considering the context and human factors that hinder or facilitate the intervention.

This study took place in primary care practices in Rio de Janeiro, which is a rich environment for clinical research due to its large dimensions, organization, and systematic use of electronic health records [97]. Most studies about telehealth have been conducted in high-income countries [29]. Hence, our findings will likely be transferable within Brazil and other countries with similar socioeconomic conditions and health care systems. Finally, we included the patients' vision on the intervention. Although provider-to-provider telehealth does not directly involve patients as participants, its ultimate goal is to improve their medical care. Patients' assessment of provider-to-provider telehealth has been investigated in a few studies by some research groups from North America [39].

## Limitations

Our trial has several limitations. The first limitation related to the study design is using a concurrent mixed methods approach where quantitative and qualitative data are collected simultaneously. This decision was driven by time and operability constraints. Nevertheless, we believe that it did not significantly affect inferences or interpretations. We relied on reports from the literature stating that concurrent designs are frequently used in health care research due to their efficiency regarding time and data collection [98].

The second limitation is the occasional synchronous communication between the primary care physicians and cardiologists during phase B, such as WhatsApp texting and audio and video calls. Although it was a deviation from the planned intervention, we decided to keep it to ensure the study's pragmatism. The interactions were not frequent, but we unfortunately did not track them as the measurement was not planned in our data collection strategy.

The third limitation is the sampling strategy for the focus groups. We had 1 focus group session with family medicine specialists and residents, 1 with patients from study phase A, and 1 with cardiologists. Of the 15 invited patients, only 5 (33%) attended the session, which could limit data availability. Therefore, a traditional data saturation assessment of the focus groups was not conducted as described in the literature [99]. Nevertheless, the researchers believe that the topics addressed in the focus groups covered most aspects of telehealth feasibility. In addition, participants mentioned other topics that enriched the content analysis. A review by Tausch and Menold [100] describes the advantages of "smaller focus group sizes for health research, especially when sensitive topics are discussed...considering 4 to 6 persons to be optimal." The aggregation of the individual interviews, originally a separate research project, further complemented the corpus of qualitative data and filled gaps by including the primary care physicians involved in phase B of the project.

The fourth limitation is that we did not include local and regional managers of primary care practices, an essential stakeholder category, as participants in this trial. As they deeply understand the work process in the practices, we may have missed crucial insights from this group. The fifth limitation concerns the study's transferability. Although the researchers assessed the sample and the corpus for analysis as satisfactory, the settings are specific to 1 practice in phase A and 1 region of Rio de Janeiro's primary care practices in phase B when considering the qualitative data collection. This may limit how the results can be generalized to other parts of the city or further geographic spaces and contexts. Regarding the quantitative methods, the large proportion of missing follow-up data undermines the outcome assessment. Therefore, all conclusions about the quantitative analysis must be seen as a trend, not a significant result. The findings are exploratory and should be interpreted cautiously. According to the CONSORT recommendations for feasibility trials and pilot studies [51], determining and attaining an adequate sample size is out of the scope of feasibility studies as the objective is not to draw statistical significance of power; otherwise, the subsequent trial would not be necessary. In any case, we relied on this result to anticipate and develop mitigation strategies for the ongoing trial, such as the active recruitment of patients based on hospital discharge lists and the inclusion of a more robust research team to ensure a higher participant recruitment success rate and better data collection [46].

## Harms and Risks

The intervention in this study inflicted minimal risk or unintended effects on the participants. However, we considered the patients' concerns about being blocked from accessing specialized consultations.

## Conclusions

Considering the described adaptations, this study showed that it is feasible to offer telehealth support from cardiologists to primary care physicians to treat patients with heart failure in

the community setting in Rio de Janeiro, Brazil. Primary care physicians found it valuable and feasible but pointed to hardships in engagement due to work overload. Patients were receptive, although they might feel unsafe if they do not have direct access to a cardiologist. Cardiologists evaluated the intervention as an attainable opportunity to connect primary and specialized care. Considering the needed modifications in recruitment and educational strategies, the intervention was assessed as suitable for the clinical trial.

## Acknowledgments

The authors would like to acknowledge primary care physicians Caio de Faria Maia and Luiz Sergio Zanini; local manager Marcelle da Silva Ribeiro from Helena Besserman Vianna primary care practice; municipality managers Amanda Aparecida Cano, Larissa Cristina Terrezo Machado, Renato Cony Seródio, and Fernanda Adães Britto; and public health researcher Raphael Mendonça Guimarães, all from Rio de Janeiro, Brazil. The Brazilian Heart Insufficiency With Telemedicine project is funded by the Danida Fellowship Centre (Danish Ministry of Foreign Affairs; start date: March 1, 2019; end date: December 31, 2024; project code: 18-M03-KU). The role of the sponsor was to provide financial support for the project, including researchers' salaries, contracts, material supply, and mobility. It did not influence the project design, data management, reporting, or publication.

## Data Availability

The datasets generated or analyzed during this study are not publicly available due to personal data protection policies but are available from the corresponding author on reasonable request.

## Authors' Contributions

LG, AFCI, MMM, VBPDF, MBD, LCMS, GPDCDS, MKG, JRLES, LPRDS, AF, and HD contributed to conceptualization. LG, ICPDN, VKF, VNM, JDSLS, and PCM contributed to data curation. LG and HD contributed to formal analysis. HD contributed to funding acquisition. LG, VKF, VNM, JDSLS, PCM, and GPDCDS contributed to investigation. LG, MKG, PCM, and HD contributed to methodology. LG, AFCI, ICPDN, and HD contributed to the project's administration. LG, AFCI, and HD contributed to resources. LG, GPDCDS, MMM, and VBPDF contributed to software. AFCI, AF, JRLES, and HD contributed to supervision. HD contributed to validation. LG and HD contributed to writing—original draft. LG, LCMS, and HD contributed to visualization. LPRDS, GPDCDS, and HD contributed to writing—review and editing. All authors have read and agreed to the published version of the manuscript.

## Conflicts of Interest

None declared.

## References

1. Reeves S, Pelone F, Harrison R, Goldman J, Zwarenstein M. Interprofessional collaboration to improve professional practice and healthcare outcomes. Cochrane Database Syst Rev 2017 Jun 22;6(6):CD000072 [FREE Full text] [doi: 10.1002/14651858.CD000072.pub3] [Medline: 28639262]
2. Bashshur R, Shannon G, Krupinski E, Grigsby J. The taxonomy of telemedicine. Telemed J E Health 2011;17(6):484-494 [FREE Full text] [doi: 10.1089/tmj.2011.0103] [Medline: 21718114]
3. WHO guideline recommendations on digital interventions for health system strengthening. World Health Organization. URL: https://apps.who.int/iris/handle/10665/357828 [accessed 2024-04-29]
4. Liddy C, Moroz I, Mihan A, Nawar N, Keely E. A systematic review of asynchronous, provider-to-provider, electronic consultation services to improve access to specialty care available worldwide. Telemed J E Health 2019 Mar;25(3):184-198 [FREE Full text] [doi: 10.1089/tmj.2018.0005] [Medline: 29927711]
5. Blank L, Baxter S, Woods HB, Goyder E, Lee A, Payne N, et al. Referral interventions from primary to specialist care: a systematic review of international evidence. Br J Gen Pract 2014 Dec 01;64(629):e765-e774 [FREE Full text] [doi: 10.3399/bjgp14x682837]
6. Smith SM, Cousins G, Clyne B, Allwright S, O'Dowd T. Shared care across the interface between primary and specialty care in management of long term conditions. Cochrane Database Syst Rev 2017 Feb 23;2(2):CD004910 [FREE Full text] [doi: 10.1002/14651858.CD004910.pub3] [Medline: 28230899]
7. Kadu MK, Stolee P. Facilitators and barriers of implementing the chronic care model in primary care: a systematic review. BMC Fam Pract 2015 Mar 06;16:12 [FREE Full text] [doi: 10.1186/s12875-014-0219-0] [Medline: 25655401]
8. Ponikowski P, Anker SD, AlHabib KF, Cowie MR, Force TL, Hu S, et al. Heart failure: preventing disease and death worldwide. ESC Heart Fail 2014 Sep;1(1):4-25 [FREE Full text] [doi: 10.1002/ehf2.12005] [Medline: 28834669]
9. Baert A, De Smedt D, De Sutter J, De Bacquer D, Puddu P, Clays E, et al. Factors associated with health-related quality of life in stable ambulatory congestive heart failure patients: systematic review. Eur J Prev Cardiol 2018 Mar;25(5):472-481 [FREE Full text] [doi: 10.1177/2047487318755795] [Medline: 29384392]

10. GBD 2019 DiseasesInjuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
11. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano G, Coats A. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res 2023 Jan 18;118(17):3272-3287 [FREE Full text] [doi: 10.1093/cvr/cvac013] [Medline: 35150240]
12. Lippi G, Sanchis-Gomar F. Global epidemiology and future trends of heart failure. AME Med J 2020 Jun;5:15 [FREE Full text] [doi: 10.21037/amj.2020.03.03]
13. Crespo-Leiro MG, Anker SD, Maggioni AP, Coats AJ, Filippatos G, Ruschitzka F. European Society of Cardiology Heart Failure Long-Term Registry (ESC-HF-LT): 1-year follow-up outcomes and differences across regions. Eur J Heart Fail 2017 Mar;19(3):438 [FREE Full text] [doi: 10.1002/ejhf.772] [Medline: 28251778]
14. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. Eur J Heart Fail 2020 Aug;22(8):1342-1356 [FREE Full text] [doi: 10.1002/ejhf.1858] [Medline: 32483830]
15. Mesquita ET, Jorge AJ, Rabelo LM, Souza Jr CV. Understanding hospitalization in patients with heart failure. Int J Cardiovasc Sci 2017;30(1):6 [FREE Full text] [doi: 10.5935/2359-4802.20160060]
16. Bragazzi NL, Zhong W, Shu J, Abu Much A, Lotan D, Grupper A, et al. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. Eur J Prev Cardiol 2021 Dec 29;28(15):1682-1690 [FREE Full text] [doi: 10.1093/eurjpc/zwaa147] [Medline: 33571994]
17. Roger VL. Epidemiology of heart failure. Circ Res 2021 May 14;128(10):1421-1434 [FREE Full text] [doi: 10.1161/circresaha.121.318172]
18. Roskvist R, Eggleton K, Arroll B, Stewart R. Non-acute heart failure management in primary care. BMJ 2024 Apr 05;385:e077057 [FREE Full text] [doi: 10.1136/bmj-2023-077057] [Medline: 38580384]
19. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q 2005 Oct 03;83(3):457-502 [FREE Full text] [doi: 10.1111/j.1468-0009.2005.00409.x] [Medline: 16202000]
20. Mendonça Guimarães R, Pimenta Ribeiro Dos Santos L, Gonçalves Pereira A, Graever L. The effect of primary care policy changes on hospitalisation for ambulatory care sensitive conditions: notes from Brazil. Public Health 2021 Dec;201:26-34 [FREE Full text] [doi: 10.1016/j.puhe.2021.09.028] [Medline: 34742114]
21. Pimenta L, Dutra VG, de Castro AL, Guimarães RM. Analysis of conditions sensitive to primary care in a successful experience of primary healthcare expansion in Brazil, 1998-2015. Public Health 2018 Sep;162:32-40 [FREE Full text] [doi: 10.1016/j.puhe.2018.05.011] [Medline: 29957336]
22. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, ESC Scientific Document Group. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021 Oct 21;42(36):3599-3726 [FREE Full text] [doi: 10.1093/eurheartj/ehab368] [Medline: 34447992]
23. Heidenreich PA, Bozkurt B, Aguilar D, Allen L, Byun J, Colvin M, ACC/AHA Joint Committee Members. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association joint committee on clinical practice guidelines. Circulation 2022 May 03;145(18):e895-1032 [FREE Full text] [doi: 10.1161/CIR.0000000000001063] [Medline: 35363499]
24. Greene SJ, Butler J, Albert NM, DeVore AD, Sharma PP, Duffy CI, et al. Medical therapy for heart failure with reduced ejection fraction: the CHAMP-HF registry. J Am Coll Cardiol 2018 Jul 24;72(4):351-366 [FREE Full text] [doi: 10.1016/j.jacc.2018.04.070] [Medline: 30025570]
25. Hayhoe B, Kim D, Aylin PP, Majeed FA, Cowie MR, Bottle A. Adherence to guidelines in management of symptoms suggestive of heart failure in primary care. Heart 2019 May;105(9):678-685 [FREE Full text] [doi: 10.1136/heartjnl-2018-313971] [Medline: 30514731]
26. Calvin JE, Shanbhag S, Avery E, Kane J, Richardson D, Powell L. Adherence to evidence-based guidelines for heart failure in physicians and their patients: lessons from the Heart Failure Adherence Retention Trial (HART). Congest Heart Fail 2012;18(2):73-78 [FREE Full text] [doi: 10.1111/j.1751-7133.2011.00263.x] [Medline: 22432552]
27. Liu XL, Wang T, Tan JY, Stewart S, Chan RJ, Eliseeva S, et al. Sustainability of healthcare professionals' adherence to clinical practice guidelines in primary care. BMC Prim Care 2022 Mar 01;23(1):36 [FREE Full text] [doi: 10.1186/s12875-022-01641-x] [Medline: 35232391]
28. van Riet EE, Hoes AW, Limburg A, Landman MA, van der Hoeven H, Rutten F. Prevalence of unrecognized heart failure in older persons with shortness of breath on exertion. Eur J Heart Fail 2014 Jul;16(7):772-777 [FREE Full text] [doi: 10.1002/ejhf.110] [Medline: 24863953]
29. Peeters KM, Reichel LA, Muris DM, Cals JW. Family physician-to-hospital specialist electronic consultation and access to hospital care: a systematic review. JAMA Netw Open 2024 Jan 02;7(1):e2351623 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.51623] [Medline: 38214930]
30. Liddy C, Drosinis P, Keely E. Electronic consultation systems: worldwide prevalence and their impact on patient care-a systematic review. Fam Pract 2016 Jun;33(3):274-285 [FREE Full text] [doi: 10.1093/fampra/cmw024] [Medline: 27075028]

31. Keely E, Liddy C, Afkham A. Utilization, benefits, and impact of an e-consultation service across diverse specialties and primary care providers. Telemed J E Health 2013 Oct;19(10):733-738 [FREE Full text] [doi: 10.1089/tmj.2013.0007] [Medline: 23980939]
32. Belber GS, Passos VD, Borysow ID, Maeyama M. Contributions of the Brazilian National Telehealth Program in the education of professionals in primary health care. Braz J Dev 2021;7(1):1198-1219 [FREE Full text] [doi: 10.34117/bjdv7n1-081]
33. Alkmim MB, Marcolino MS, Figueira RM, Sousa L, Nunes MS, Cardoso CS, et al. Factors associated with the use of a teleconsultation system in Brazilian primary care. Telemed J E Health 2015 Jun;21(6):473-483. [doi: 10.1089/tmj.2014.0112] [Medline: 25785650]
34. Maldonado JM, Marques AB, Cruz A. Telemedicine: challenges to dissemination in Brazil. Cad Saude Publica 2016 Nov 03;32Suppl 2(Suppl 2):e00155615 [FREE Full text] [doi: 10.1590/0102-311X00155615] [Medline: 27828681]
35. Roman R, Lima KM, Fontoura Moreira A, Umpierre RN, Hauser L, Rados D. Distance education to improve the quality of asthma treatment in primary health care: cluster randomized clinical trial. Rev Bras Med Fam Community 2019;14(41):2065 [FREE Full text] [doi: 10.5712/rbmfc14(41)2065]
36. Schmitz CA, Harzheim E. Oferta e utilização de teleconsultorias para Atenção Primária à Saúde no Programa Telessaúde Brasil Redes. Rev Bras Med Fam 2017;12(39):1-11 [FREE Full text] [doi: 10.5712/rbmfc12(39)1453]
37. Soriano Marcolino M, Minelli Figueira R, Pereira Afonso Dos Santos J, Silva Cardoso C, Luiz Ribeiro A, Alkmim M. The experience of a sustainable large scale Brazilian telehealth network. Telemed J E Health 2016 Nov;22(11):899-908 [FREE Full text] [doi: 10.1089/tmj.2015.0234] [Medline: 27167901]
38. Haleem A, Javaid M, Singh RP, Suman R. Telemedicine for healthcare: capabilities, features, barriers, and applications. Sens Int 2021;2:100117 [FREE Full text] [doi: 10.1016/j.sintl.2021.100117] [Medline: 34806053]
39. Osman MA, Schick-Makaroff K, Thompson S, Bialy L, Featherstone R, Kurzawa J, et al. Barriers and facilitators for implementation of electronic consultations (eConsult) to enhance access to specialist care: a scoping review. BMJ Glob Health 2019;4(5):e001629 [FREE Full text] [doi: 10.1136/bmjgh-2019-001629] [Medline: 31565409]
40. Kozlowska O, Lumb A, Tan GD, Rea R. Barriers and facilitators to integrating primary and specialist healthcare in the United Kingdom: a narrative literature review. Future Healthc J 2018 Mar;5(1):64-80 [FREE Full text] [doi: 10.7861/futurehosp.5-1-64] [Medline: 31098535]
41. Robert G, Greenhalgh T, MacFarlane F, Peacock R. Adopting and assimilating new non-pharmaceutical technologies into health care: a systematic review. J Health Serv Res Policy 2010 Oct;15(4):243-250 [FREE Full text] [doi: 10.1258/jhsrp.2010.009137] [Medline: 20592046]
42. Greenhalgh T, Papoutsi C. Spreading and scaling up innovation and improvement. BMJ 2019 May 10;365:l2068 [FREE Full text] [doi: 10.1136/bmj.l2068] [Medline: 31076440]
43. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004 Dec 09;82(4):581-629 [FREE Full text] [doi: 10.1111/j.0887-378x.2004.00325.x]
44. Gonçalves-Bradley DC, J Maria AR, Ricci-Cabello I, Villanueva G, Fønhus MS, Glenton C, et al. Mobile technologies to support healthcare provider to healthcare provider communication and management of care. Cochrane Database Syst Rev 2020 Aug 18;8(8):CD012927 [FREE Full text] [doi: 10.1002/14651858.CD012927.pub2] [Medline: 32813281]
45. Vimalananda VG, Orlander JD, Afable MK, Fincke BG, Solch AK, Rinne ST, et al. Electronic consultations (E-consults) and their outcomes: a systematic review. J Am Med Inform Assoc 2020 Mar 01;27(3):471-479 [FREE Full text] [doi: 10.1093/jamia/ocz185] [Medline: 31621847]
46. Graever L, Issa AF, Fonseca VB, Melo MM, Silva GP, Nóbrega IC, et al. Telemedicine support for primary care providers versus usual care in patients with heart failure: protocol of a pragmatic cluster randomised trial within the Brazilian Heart Insufficiency with Telemedicine (BRAHIT) study. Int J Environ Res Public Health 2023 May 24;20(11):5933 [FREE Full text] [doi: 10.3390/ijerph20115933] [Medline: 37297537]
47. Brazilian Heart Insufficiency with Telemedicine (BRAHIT). Department of Biomedical Sciences, University of Copenhagen. 2019. URL: https://bmi.ku.dk/english/brahit [accessed 2024-04-29]
48. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Oct 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
49. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner K, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
50. Bardin L. L'analyse de contenu. Paris, France: Presses Universitaires de France; 2013.
51. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239 [FREE Full text] [doi: 10.1136/bmj.i5239] [Medline: 27777223]

52. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Int J Surg 2014 Dec;12(12):1500-1524 [FREE Full text] [doi: 10.1016/j.ijsu.2014.07.014] [Medline: 25046751]
53. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Oct;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
54. Braun V, Clarke V. Reporting guidelines for qualitative research: a values-based approach. Qual Res Psychol 2024 Oct 30;22(2):399-438 [FREE Full text] [doi: 10.1080/14780887.2024.2382244]
55. Watkins DC, Gioia D. “Sixth floor”: writing up, presenting, and teaching mixed methods. In: Watkins D, Gioia D, editors. Mixed Methods Research. Oxford, UK: Oxford Academic Press; 2015:139-170.
56. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5:114 [FREE Full text] [doi: 10.1186/s40814-019-0499-1] [Medline: 31608150]
57. O'Cathain A, Hoddinott P, Lewin S, Thomas KJ, Young B, Adamson J, et al. Maximising the impact of qualitative research in feasibility studies for randomised controlled trials: guidance for researchers. Pilot Feasibility Stud 2015;1:32 [FREE Full text] [doi: 10.1186/s40814-015-0026-y] [Medline: 27965810]
58. Macinko J, Harris M. Brazil's family health strategy. N Engl J Med 2015 Sep 24;373(13):1277-1278. [doi: 10.1056/nejmc1509056]
59. Soranz D, Pinto LF, Penna GO. Themes and Reform of Primary Health Care (RCAPS) in the city of Rio de Janeiro, Brazil. Cien Saude Colet 2016 May;21(5):1327-1338 [FREE Full text] [doi: 10.1590/1413-81232015215.01022016] [Medline: 27166884]
60. Liddy C, Moroz I, Afkham A, Keely E. Sustainability of a primary care-driven eConsult service. Ann Fam Med 2018 Mar 12;16(2):120-126 [FREE Full text] [doi: 10.1370/afm.2177] [Medline: 29531102]
61. Transkriptor: convert audio or video to text. Transkriptor. URL: https://app.transkriptor.com/dashboard [accessed 2024-04-29]
62. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383 [FREE Full text] [doi: 10.3758/bf03195514] [Medline: 14587545]
63. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 May;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
64. Bazeley P, Jackson K. Qualitative Data Analysis with NVivo. 2nd edition. Thousand Oaks, CA: SAGE Publications; 2013.
65. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 May;62(1):107-115 [FREE Full text] [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
66. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Mar;24(2):105-112 [FREE Full text] [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
67. Jones K, Voutsina C. Qualitative content analysis: theoretical background and procedures. In: Bikner-Ahsbahs A, Knipping C, Presmeg N, editors. Approaches to Qualitative Research in Mathematics Education: Examples of Methodology and Methods. Cham, Switzerland: Springer; Aug 31, 2017:365-390.
68. Rohde LE, Montera MW, Bocchi EA, Clausell NO, Albuquerque DC, Rassi S, et al. Diretriz Brasileira de Insuficiência Cardíaca Crônica e Aguda. Arq Bras Cardiol 2018 Oct;111(3):436-539 [FREE Full text] [doi: 10.5935/abc.20180190] [Medline: 30379264]
69. Patel J, Rassekh N, Fonarow GC, Deedwania P, Sheikh F, Ahmed A, et al. Guideline-directed medical therapy for the treatment of heart failure with reduced ejection fraction. Drugs 2023 Jul;83(9):747-759 [FREE Full text] [doi: 10.1007/s40265-023-01887-4] [Medline: 37254024]
70. Aschbrenner KA, Kruse G, Gallo JJ, Plano Clark VL. Applying mixed methods to pilot feasibility studies to inform intervention trials. Pilot Feasibility Stud 2022 Oct 26;8(1):217 [FREE Full text] [doi: 10.1186/s40814-022-01178-x] [Medline: 36163045]
71. Bugge C, Williams B, Hagen S, Logan J, Glazener C, Pringle S, et al. A process for Decision-making after Pilot and feasibility Trials (ADePT): development following a feasibility study of a complex intervention for pelvic organ prolapse. Trials 2013 Oct 25;14:353 [FREE Full text] [doi: 10.1186/1745-6215-14-353] [Medline: 24160371]
72. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
73. Hung CL, Chao TF, Su CH, Liao JN, Sung KT, Yeh HI, et al. Income level and outcomes in patients with heart failure with universal health coverage. Heart 2021 Mar;107(3):208-216 [FREE Full text] [doi: 10.1136/heartjnl-2020-316793] [Medline: 33082175]
74. Dantas MN, de Souza DL, de Souza AM, Aiquoc KM, Souza TA, Barbosa I. Factors associated with poor access to health services in Brazil. Rev Bras Epidemiol 2020;24:e210004 [FREE Full text] [doi: 10.1590/1980-549720210004] [Medline: 33331413]

75. Ware J, Mawby R. Patient access to general practice: ideas and challenges from the front line. Royal College of General Practitioners. URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=92b60ee0a70d2d54a92208d56a2a8fa64cec662c [accessed 2024-04-29]
76. Foy R, Hempel S, Rubenstein L, Suttorp M, Seelig M, Shanman R, et al. Meta-analysis: effect of interactive communication between collaborating primary care physicians and specialists. Ann Intern Med 2010 Mar 16;152(4):247-258 [FREE Full text] [doi: 10.7326/0003-4819-152-4-201002160-00010] [Medline: 20157139]
77. Liddy C, Maranger J, Afkham A, Keely E. Ten steps to establishing an e-consultation service to improve access to specialist care. Telemed J E Health 2013 Dec;19(12):982-990 [FREE Full text] [doi: 10.1089/tmj.2013.0056] [Medline: 24073898]
78. Doniec K, Dall'Alba R, King L. Austerity threatens universal health coverage in Brazil. Lancet 2016 Aug 27;388(10047):867-868 [FREE Full text] [doi: 10.1016/S0140-6736(16)31428-3] [Medline: 27597461]
79. Melo EA, de Mendonça MH, Teixeira M. The economic crisis and primary health care in the SUS of Rio de Janeiro, Brazil. Cien Saude Colet 2019 Dec;24(12):4593-4598 [FREE Full text] [doi: 10.1590/1413-812320182412.25432019] [Medline: 31778509]
80. O'Dwyer G, Graever L, Britto F, Menezes T, Konder M. Financial crisis and healthcare: the case of the municipality of Rio de Janeiro, Brazil. Cien Saude Colet 2019 Dec;24(12):4555-4568 [FREE Full text] [doi: 10.1590/1413-812320182412.23212019] [Medline: 31778505]
81. Juliani C, MacPhee M, Spiri W. Brazilian specialists' perspectives on the patient referral process. Healthcare (Basel) 2017 Jan 29;5(1):4 [FREE Full text] [doi: 10.3390/healthcare5010004] [Medline: 28146046]
82. Liddy C, Moroz I, Keely E, Taljaard M, Deri Armstrong C, Afkham A, et al. Understanding the impact of a multispecialty electronic consultation service on family physician referral rates to specialists: a randomized controlled trial using health administrative data. Trials 2019 Jul 10;20(1):348 [FREE Full text] [doi: 10.1186/s13063-019-3393-5] [Medline: 31182123]
83. Haroon SM, Jordan RE, O'Beirne-Elliman J, Adab P. Effectiveness of case finding strategies for COPD in primary care: a systematic review and meta-analysis. NPJ Prim Care Respir Med 2015 Aug 27;25(1):15056 [FREE Full text] [doi: 10.1038/npjpcrm.2015.56] [Medline: 26313400]
84. Lanham HJ, Leykum LK, Taylor BS, McCannon CJ, Lindberg C, Lester R. How complexity science can inform scale-up and spread in health care: understanding the role of self-organization in variation across local contexts. Soc Sci Med 2013 Oct;93:194-202 [FREE Full text] [doi: 10.1016/j.socscimed.2012.05.040] [Medline: 22819737]
85. Greenhalgh T, Russell J. Why do evaluations of eHealth programs fail? An alternative set of guiding principles. PLoS Med 2010 Dec 02;7(11):e1000360 [FREE Full text] [doi: 10.1371/journal.pmed.1000360] [Medline: 21072245]
86. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, Van Spall HG. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74 [FREE Full text] [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]
87. Knox M, Murphy EJ, Leslie T, Wick R, Tuot DS. e-Consult implementation success: lessons from 5 county-based delivery systems. Am J Manag Care 2020 Jan 01;26(1):e21-e27 [FREE Full text] [doi: 10.37765/ajmc.2020.42149] [Medline: 31951363]
88. Hassan Alshehri HM, Alshehri A, Alshehri A, Alshehri S, Al Zahib YH, Alahmari A, et al. Primary health care professionals' opinions regarding continuing medical education: a cross sectional study. Medicine (Baltimore) 2024 Dec 13;103(50):e40865 [FREE Full text] [doi: 10.1097/MD.0000000000040865] [Medline: 39686415]
89. Easley J, Wassersug R, Matthias S, Tompson M, Schneider ND, O'Brien MA, et al. Patient engagement in health research: perspectives from patient participants. Curr Oncol 2023 Mar 26;30(3):2770-2780 [FREE Full text] [doi: 10.3390/curroncol30030210] [Medline: 36975423]
90. DE Albuquerque DC, DE Barros E Silva PG, Lopes R, Hoffmann-Filho C, Nogueira P, Reis H, BREATHE INVESTIGATORS. In-hospital management and long-term clinical outcomes and adherence in patients with acute decompensated heart failure: primary results of the first Brazilian Registry of Heart Failure (BREATHE). J Card Fail 2024 May;30(5):639-650 [FREE Full text] [doi: 10.1016/j.cardfail.2023.08.014] [Medline: 37648061]
91. de Albuquerque NL, de Oliveira FJ, Machado LD, de Araujo TL, Caetano J, Aquino PD. Social determinants of health and heart failure hospitalizations in Brazil. Rev Esc Enferm 2020:54 [FREE Full text]
92. Kronenfeld JJ. Impact of Demographics on Health and Healthcare: Race, Ethnicity and Other Social Factors. Rio de Janeiro, Brazil: Emerald Publishing; 2010.
93. Oliveira GM, Brant LC, Polanczyk CA, Biolo A, Nascimento BR, Malta DC, et al. Cardiovascular statistics – Brazil. Arq Bras Cardiol 2020;115(3):308-439 [FREE Full text]
94. Chauhan A, Walton M, Manias E, Walpola R, Seale H, Latanik M, et al. The safety of health care for ethnic minority patients: a systematic review. Int J Equity Health 2020 Jul 08;19(1):118 [FREE Full text] [doi: 10.1186/s12939-020-01223-2] [Medline: 32641040]
95. Erves JC, Mayo-Gamble TL, Malin-Fair A, Boyer A, Joosten Y, Vaughn Y, et al. Needs, priorities, and recommendations for engaging underrepresented populations in clinical research: a community perspective. J Community Health 2017 Jul;42(3):472-480 [FREE Full text] [doi: 10.1007/s10900-016-0279-2] [Medline: 27812847]
96. Packer M. Is Any Patient with Chronic Heart Failure Receiving the Right Dose of the Right Beta-Blocker in Primary care? Am J Med 2019 Dec;132(11):e761-e762 [FREE Full text] [doi: 10.1016/j.amjmed.2019.03.026] [Medline: 30986395]

97. Coeli CM, Saraceni V, Medeiros PM, da Silva Santos HP, Guillen L, Alves LG, et al. Record linkage under suboptimal conditions for data-intensive evaluation of primary care in Rio de Janeiro, Brazil. BMC Med Inform Decis Mak 2021 Jul 15;21(1):190 [FREE Full text] [doi: 10.1186/s12911-021-01550-6] [Medline: 34130670]
98. Meissner H, Creswell J, Klassen AC, Plano V, Smith KC. Best practices for mixed methods research in the health sciences. Office of Behavioral and Social Sciences Research. URL: https://obssr.od.nih.gov/sites/obssr/files/Best_Practices_for_Mixed_Methods_Research.pdf [accessed 2024-04-29]
99. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
100. Tausch AP, Menold N. Methodological aspects of focus groups in health research: results of qualitative interviews with focus group moderators. Glob Qual Nurs Res 2016;3:2333393616630466 [FREE Full text] [doi: 10.1177/2333393616630466] [Medline: 28462326]

## Abbreviations

**BRAHIT:** Brazilian Heart Insufficiency with Telemedicine
**CONSORT:** Consolidated Standards of Reporting Trials
**GDMT:** guideline-directed medical therapy
**HFrEF:** heart failure with reduced ejection fraction
**INC:** National Institute of Cardiology
**RAAS-I:** renin-angiotensin-aldosterone system inhibitors
**REDCap:** Research Electronic Data Capture
**SGLT-2:** sodium-glucose cotransporter 2

*Edited by A Coristine; submitted 23.07.24; peer-reviewed by T Manavi, F Epelde; comments to author 08.01.25; revised version received 10.02.25; accepted 10.03.25; published 17.04.25.*

*Please cite as:*
*Graever L, Mafra PC, Figueira VK, Miler VN, Sobreiro JDSL, Silva GPDCD, Issa AFC, Savassi LCM, Dias MB, Melo MM, Fonseca VBPD, Nóbrega ICPD, Gomes MK, Santos LPRD, Lapa e Silva JR, Froelich A, Dominguez H*
*Telehealth Support From Cardiologists to Primary Care Physicians in Heart Failure Treatment: Mixed Methods Feasibility Study of the Brazilian Heart Insufficiency With Telemedicine Trial*
*JMIR Cardio 2025;9:e64438*
*URL: https://cardio.jmir.org/2025/1/e64438*
*doi:10.2196/64438*
*PMID:*

# Patient Perspectives on the "Future Patient" Telerehabilitation Program for Atrial Fibrillation: Qualitative Study

Elisabet Dortea Ragnvaldsdóttir Joensen[1], BSc; Andi Eie Albertsen[2], MD, PhD; Helle Spindler[3], MSc, PhD; Katja Møller Jensen[1], MSc; Lars Frost[4], MD, PhD; Lars Dittmann[5], MSc, PhD; Mathushan Gunasegaram[1], MSc; Søren Paaske Johnsen[6], MD, PhD; Mads Rovsing Jochumsen[1], MSc, PhD; Dorthe Svenstrup[2], MD, PhD; Birthe Dinesen[1], MSc, PhD

[1]Department of Health Science and Technology, Aalborg University, Selma Lagerløfsvej 249, Aalborg, Denmark

[2]Regionshospitalet Viborg, Viborg, Denmark

[3]Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus, Denmark

[4]Regionshospitalet Silkeborg, Silkeborg, Denmark

[5]Department of Electrical and Photonics Engineering, Danish Technical University, Copenhagen, Denmark

[6]Danish Center for Health Services Research, Aalborg University, Aalborg, Denmark

**Corresponding Author:**
Elisabet Dortea Ragnvaldsdóttir Joensen, BSc
Department of Health Science and Technology, Aalborg University, Selma Lagerløfsvej 249, Aalborg, Denmark

## Abstract

**Background:** Atrial fibrillation (AF) is a prevalent chronic condition with increasing incidence worldwide. AF increases the risks of stroke, heart failure, and myocardial infarction and imposes a substantial burden on the health care system. Cardiac rehabilitation programs, while effective, often have low patient adherence. Recent evidence suggests that cardiac telerehabilitation, where patients are given home monitoring devices, could enhance adherence and outcomes. The program "Future Patient—Telerehabilitation of Patients with AF" (FP-AF) was created to assess the effects and potential benefits of cardiac telerehabilitation on patients with AF.

**Objective:** The objective of this study is to explore the experiences of patients participating in the FP-AF program.

**Methods:** This qualitative sub-study is part of the multicenter, randomized controlled FP-AF trial, which included 208 patients. Semi-structured interviews were conducted on 14 patients, randomly selected from participants in the intervention arm of the FP-AF program. The patient interviews, guided by self-determination theory, focused on patients' experiences with the FP-AF program, including the use of telerehabilitation technologies and a web-based portal called the "HeartPortal." Interview responses were analyzed using NVivo software (version 14.0; QSR International), with thematic coding based on interview guides and methodological guidance elaborated by Brinkmann & Kvale. The study adhered to ethical guidelines, with informed consent obtained from all participants.

**Results:** Based on the interviews, the following themes were identified: the home monitoring devices are viewed positively by the patients; the HeartPortal is a useful digital toolbox; patients develop new coping strategies for living with AF; the measured values are useful for the patients; the community of practice is beneficial; and the FP-AF program creates a sense of security.

**Conclusions:** Participation in the FP-AF program enhanced patients' sense of security, empowerment, and knowledge about AF. This improvement was due largely to a combination of patients' use of the HeartPortal and the educational sessions at health care centers. Telerehabilitation for patients with AF may be a useful way of researching this group of patients with a focus on rehabilitation and may be an effective means of offering rehabilitation to this group in the future.

**Trial Registration:** ClinicalTrials.gov NCT06101485; https://clinicaltrials.gov/study/NCT06101485

**International Registered Report Identifier (IRRID):** RR2-10.2196/64259

***(JMIR Cardio 2025;9:e68663)*** doi:10.2196/68663

**KEYWORDS**

atrial fibrillation; telerehabilitation; qualitative interviews; patient education; home monitoring

## Introduction

Atrial fibrillation (AF) is a chronic cardiovascular condition with a lifetime risk affecting about 1 out of 3-5 individuals aged 45 years or older, depending on their risk factor profile [1-3]. Risk factors for AF include weight, hypertension, physical activity, diet, alcohol consumption, and smoking status, as well as comorbidities such as type 2 diabetes mellitus, sleep apnea, heart failure, and myocardial infarction [1,3]. The incidence of AF is growing due to an aging population, improved opportunistic screening for asymptomatic AF, and an increase in modifiable risk factors [1,2,4]. Untreated AF is associated with significant risks, including a fivefold increase in the risk of stroke and heart failure, as well as a twofold increase in the risk of myocardial infarction and excess mortality [1,5]. In addition, AF imposes a substantial economic burden on health care systems [6].

AF is known to negatively affect quality of life (QoL) and restrict patients' ability to carry out daily activities [7]. However, these negative impacts may be mitigated through patient empowerment [8]. The World Health Organization defines empowerment as "a process through which people gain greater control over decisions and actions affecting their health" [9]. Health care professionals can promote patient empowerment by implementing a patient-centered approach [10,11], with a key focus on enhancing patients' understanding of their condition [11]. Patients with AF may benefit from cardiac rehabilitation (CR) programs specifically tailored to help them manage and live with their condition [12]. A Cochrane review highlights that exercise-based rehabilitation for patients with AF reduces symptoms and recurrence and improves QoL and exercise capacity [13].

CR includes health management interventions that provide patients with the necessary knowledge and support to manage their disease through patient education, exercise, risk-management strategies, and psychological support [12]. A systematic review of educational interventions for patients with AF states that patient education is associated with a decrease in mortality and readmission, as well as having a positive impact on psychological factors such as anxiety, depression, and QoL [14]. Despite these benefits, studies have shown that adherence to CR is low. Factors associated with lower adherence to CR include female gender, older age, unemployment, comorbidities, and geographical barriers [15,16].

A Cochrane review [17] comparing home-based and center-based CR found that both types of CR were similar in their effects on QoL, modifiable risk factors, exercise capacity, mortality, and hospital admission. The Cochrane review found a small but significantly higher completion rate for home-based CR compared to the center-based CR [17].

A recent literature review by Owen and O'Carroll [18] found that cardiac telerehabilitation (CTR) had a level of effectiveness equal to that of center-based CR in outcomes such as physical activity, weight, blood pressure, QoL, depression, and anxiety. Additionally, the telerehabilitation groups' adherence to the rehabilitation program was higher than center-based rehabilitation [18]. In another study, Cai et al [19] found that patients with AF undergoing CTR had significantly increased cardiac capacity compared to those in conventional CR. Similar findings are reported by Pagliari et al [20], who found that CTR led to significantly increased exercise capacity. Furthermore, Cai et al [19] found significant improvements in health beliefs and physical activity in both groups. These findings indicate that CTR, which takes place in the patient's home environment, could be a suitable alternative to CR because it would generate a potentially higher level of adherence to rehabilitation programs. CTR may either supplement or serve as an alternative to center-based CR. A CTR solution delivers one or more rehabilitation modules directly to the patient through technologies such as wearables, smartphones, and video calls, allowing patients to participate from their own homes [21]. In order to assess the effects of CTR on patients with AF, the "Future Patient" program has been developed.

The educational CTR program "Future Patient—Telerehabilitation of Patients With AF" was developed through a co-creation process involving patients with AF, their relatives, and researchers. The program was evaluated in a pilot study by Dinesen et al [22], where it was found to be useful by patients with AF and their relatives. In particular, our CTR program was found to enhance patients' sense of security, increase their knowledge about symptom management, and promote a community of practice that connected patients and their relatives with health care professionals. Following the pilot study, the FP-AF program is now being evaluated in a multicenter, mixed-methods, randomized controlled trial, which began enrolling patients in January 2023 and is expected to conclude in June 2025.

This study aimed to explore the experiences of patients with AF participating in the FP-AF program.

## Methods

### Qualitative Study

The present study is a qualitative sub-study within the multicenter, mixed-methods randomized controlled trial on the FP-AF program (FP-AF study), which includes a total of 208 patients [23]. This qualitative sub-study uses a triangulation of data collection techniques: document analysis, patient observation, and semi-structured interviews with patients from the intervention group. Furthermore, a user panel was established, with meetings held twice a year during the FP-AF study. This study was reported in accordance with the Consolidated Criteria for Reporting Qualitative Research checklist [24].

### Participants

Patients diagnosed with AF at the Departments of Cardiology at Silkeborg, Viborg, and Skive Regional Hospitals (Denmark) were assessed for eligibility for the FP-AF study. The inclusion criteria were as follows: the patient must be diagnosed with AF; be an adult aged 18 years or older; live in the Skive, Viborg, or Silkeborg municipalities; live at home and be capable of caring for themselves; and possess basic computer skills or have a relative or friend with basic computer skills. The exclusion criteria were as follows: pregnancy; refusal or inability to

cooperate; patients who did not speak, read, or understand Danish; and patients with a life expectancy of less than a year, based on clinical judgment and underlying medical conditions.

In selecting the interviewed patients, we used a random selection process, ensured equal distribution of females and males, and selected those patients who had participated in the patient education module at the health care centers. The participants were contacted by telephone by a research assistant and invited to participate in the interviews. In total, 18 randomly selected patients were contacted, of whom 3 were unable to participate on the suggested dates and 1 did not wish to participate. A total of 14 patients were interviewed. The patients have not recieved any compensation for their participation in the interviews.

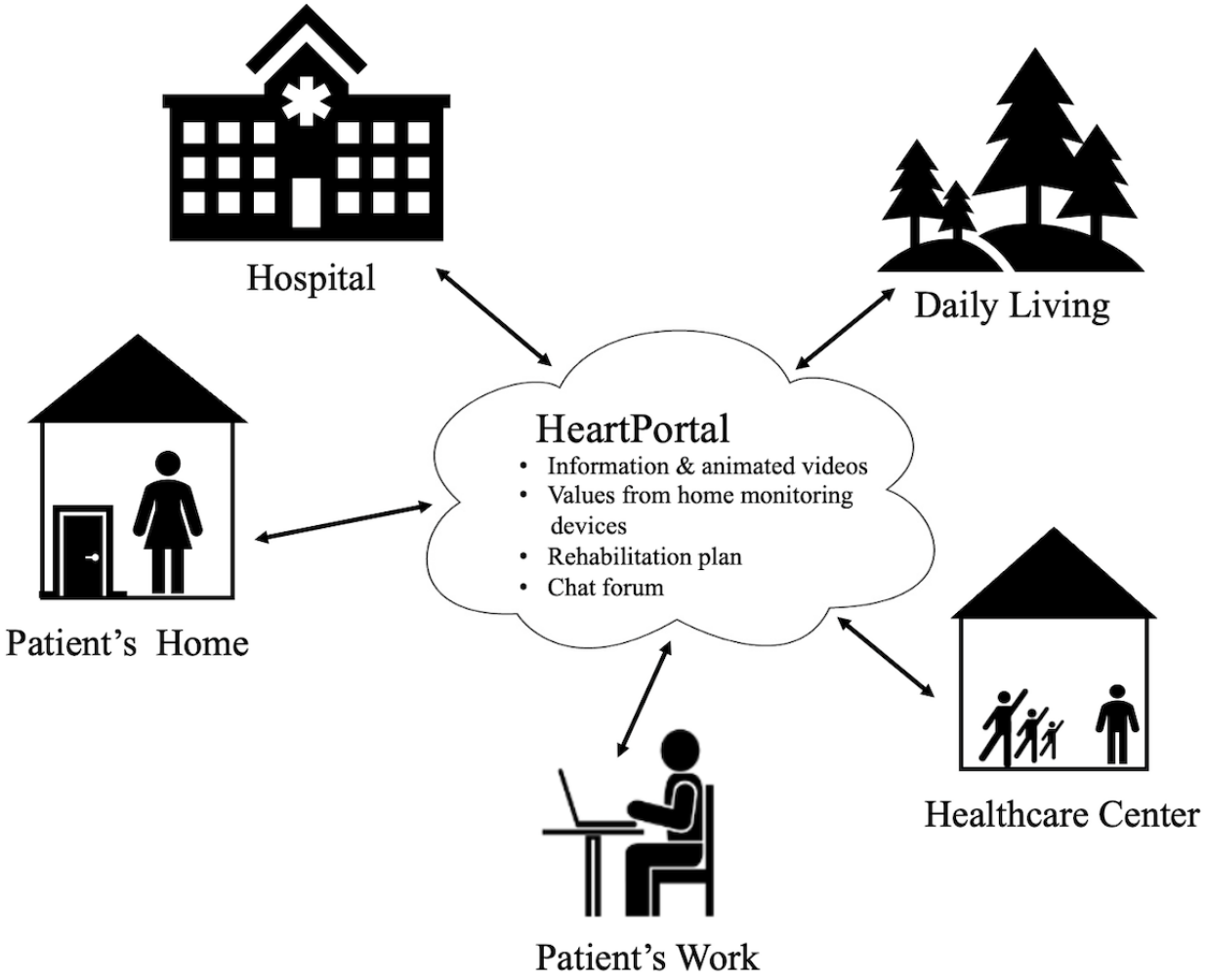
### Ethical Considerations

The FP-AF study was conducted in accordance with the Helsinki Declaration, and all participants were asked to sign an informed consent form before entering the study. Upon enrollment in the FP-AF study, all participants agreed to participate in interviews regarding their experiences with the study. The FP-AF study has been approved by the regional Ethics Committee (N-20220056) and is listed in ClinicalTrials.gov (NCT06101485). Participants were informed that they have the right to withdraw their consent at any time during the study, and the reason for their withdrawal will also be documented if participants so wish. An agreement on data sharing has been established between the participants and the researchers.

### Interventions

The FP-AF program is targeted at both patients with AF and their relatives. The program consists of two modules: (1) an education and monitoring module using telerehabilitation technologies, lasting 4 months and (2) a follow-up module, where the patients use their own personal devices to measure steps and have access to the HeartPortal, lasting 3 months. Upon enrollment, patients in the intervention group were given several home monitoring devices, including a blood pressure monitor, weight scale, activity tracker, electrocardiogram (ECG) monitor, and sleep sensor. All patients had individual meetings with the project nurses, during which they received instruction on how to use the technologies. The HeartPortal is a web-based portal accessible to patients and health care professionals. Project nurses review the patients' measurements twice a week and have continuing contact with the patients through the HeartPortal. Patients could give their consent to allow their relatives to be able to access the HeartPortal. In addition, patients and their relatives were invited to participate in an educational course at the local health care center, covering themes such as living with AF, recognizing symptoms of AF, managing the disease using digital technologies, and supporting a spouse or partner with AF. Patients in the control group received the conventional AF education delivered in person at the hospital. The education program consists of 4 sessions, each lasting 3 hours. The context of the FP-AF program can be seen in Figure 1.

**Figure 1.** Context of the FP-AF program. FP-AF: Future Patient—Telerehabilitation of Patients with AF.




### The HeartPortal

The HeartPortal was developed through a participatory design process and functions as both a digital toolbox and a web-based learning module. Patients and their relatives in the intervention group have access to the HeartPortal, where they can receive information about AF through text and animated short films. They can also view visualizations of their measured data (blood pressure, weight, steps, and sleep) and communicate via chat and video with health care professionals at the hospital and health care centers. The data from the ECG can be accessed on a separate platform.

### Theoretical Approach: Self-Determination Theory

The FP-AF program is based on self-determination theory (SDT). SDT highlights motivation as an essential component of any successful rehabilitation [25]. In this light, SDT identifies 3 basic needs for human motivation: (1) autonomy, meaning that the patient identifies with the goals of rehabilitation and values these goals as personally important; (2) competence, meaning that patients believe themselves to have (acquired) the necessary skills and knowledge to achieve their goals and receive appropriate feedback and guidance; and (3) relatedness, meaning that the health care professionals and social network create an environment where the patient feels supported, respected, and understood. To ensure continuous engagement, the motivation must be intrinsic to the patient, such that all 3 primary needs must be supported simultaneously [25].

### Document Analysis and Patient Observation

Document analysis was conducted through a systematic evaluation of written materials, including homepages on digital health strategies, rehabilitation policies, and homepages from the involved health care organizations; these materials enabled us to assemble relevant background knowledge for the context of this study. The patient observation was conducted during the education modules held at the health care centers, and the aim here was to observe the quality and level of patient engagement.

### Semi-Structured Interviews

Semi-structured exploratory interviews, inspired by Brinkmann and Kvale [26], were conducted with 14 patients from the intervention group in the FP-AF study. These interviews took place upon the completion of the program. An interview guide was designed in several steps: (1) Background questions were formulated in order to start the interview, eg, the patient is asked to present themselves and their disease; (2) based on the theoretical framework, key themes were extracted for the interview guide, eg, motivation, competences, and autonomy; (3) a total of 4 interviews were transcribed, read, and coded by 2 researchers individually. After the coding, the researchers compared the codes to ensure that they shared the same

understanding of the codes and to ensure intersubjectivity. Via this process, new themes and questions were identified for the interview guide, eg, on the importance of participation in education at the local health care center; and (4) finally, the interview guide was pilot tested, adjusted, and prepared for use. The interview guide covered the following themes: the use of technologies and the HeartPortal, the patients' experiences with the telerehabilitation program, their AF knowledge, competencies, communication with health care professionals, and the perceived advantages or disadvantages of the program.

The interviews were conducted by the first (EDRJ, BSc, student assistant) and last (BD, PhD, professor) female authors, who have years of experience conducting qualitative research. The interviewers had no prior personal relationships with the patients. The interviews took place in the patients' homes and lasted 30 - 60 minutes. All interviews were tape-recorded, and the interviewers took notes throughout. Condensed data were presented for a user panel in the FP-AF program. After 10 interviews, the research team agreed that they had reached a point of data saturation, defined as the point where no new themes or information emerged. The aim of the user panel has been to have a panel of patients and their relatives who can give second opinions and feedback on relevant issues in the project. The panel consisted of 9 patients who had participated in the FP-AF program, with new patients joining the panel as the study progressed. Preliminary findings from the interviews were discussed at the user panel meeting in May 2024 in order to validate whether data saturation had been reached. After the discussion, the patients advised the research team to conduct more interviews. After 4 more interviews, the research team found that no new themes or information emerged and concluded that data saturation was reached.

## Analysis of Data

A research assistant transcribed all interviews into text files using Microsoft Word. Any identifying information regarding the patients was removed from the transcript, and their names were anonymized with alias names. The transcribed interviews were then coded in NVivo (version 14.0; QSR International), guided by the SDT theoretical framework and by interview analysis methodologies developed by Brinkmann and Kvale [26]. Two researchers developed the code tree, defining overall themes and subthemes based on the interview guide, concepts derived from the theoretical framework, and key findings from the analysis of 3 random interviews. Before the analysis of the interviews, the code tree was reviewed and discussed by the 2 researchers to ensure intersubjectivity. The code tree was then used for the analysis of the interviews with a focus on the experiences of patients with AF participating in the FP-AF study. The data have been condensed and are presented in Table 1.

**Table .** Identified themes and subthemes from interviewed patients in the FP-AF[a] study.

| Themes | Subthemes |
|---|---|
| Devices have many functions | • User-friendly technology (n=11)<br>• ECG[b] creates a sense of security (n=3)<br>• Step counter motivates me to exercise (n=5) |
| HeartPortal as a digital toolbox | • Functions like a toolbox to navigate AF (n=5)<br>• Used as a communication platform with health care professionals (n=8)<br>• Have not used the HeartPortal frequently (n=5) |
| Coping strategies for living with AF | • Increased knowledge to handle own symptoms (n=12)<br>• Feel empowered to handle own disease (n=4) |
| Measured values are useful | • Give an overview of own data (n=5)<br>• Creates a sense of security (n=6)<br>• Careful not to be overly concerned with measuring my values (n=2) |
| Community of practice | • Community of practice with peers is beneficial (n=10)<br>• Experienced patients with AF already have the knowledge (n=2)<br>• Meeting and talking to other patients can increase your awareness negatively (n=1) |
| FP-AF program creates a sense of security | • Information provided by the health care center is useful (n=10)<br>• The patient education creates a sense of security (n=11)<br>• From a sense of unease to a sense of calm (n=2)<br>• Being monitored at home is beneficial (n=2)<br>• No need for education at health care center (n=2) |

[a]FP-AF: Future Patient—Telerehabilitation of Patients with Atrial Fibrillation.

[b]ECG: electrocardiogram.

## Results

The interviewed patients ranged in age from 58 to 93 (mean 70 [SD 8] years). There was an equal distribution of men and women in the interviews, and the patients were predominantly diagnosed with paroxysmal AF (n=11, 79%), married (n=10, 71%), had a vocational qualification (n=9, 64%), and were largely retired (n=10, 71%). Table 2 shows the baseline characteristics of the patients.

**Table .** Baseline characteristics of patients interviewed (N=14) in the FP-AF[a] study.

| Variable | Value |
|---|---|
| Gender, n (%) | |
| Males | 7 (50) |
| Females | 7 (50) |
| Age (years), mean (SD) | |
| Males | 70.4 (11.0) |
| Females | 70 (5.3) |
| Primary diagnosis, n (%) | |
| Paroxysmal | 11 (78.6) |
| Persistent | 2 (14.3) |
| Permanent | 1 (7.1) |
| Clinical parameters, mean (SD) | |
| Weight (kg) | 82 (16.5) |
| Height (cm) | 171.6 (9.7) |
| Systolic blood pressure (mmHg) | 127.6 (12.5) |
| Diastolic blood pressure (mmHg) | 84.1 (8.3) |
| Pulse (beats/min) | 62.6 (10.4) |
| Ejection fraction (%) | 58.2 (6.7) |
| $CHA_2DS_2VASc$ score | 1.9 (1.4) |
| Secondary diagnosis, n (%) | |
| Heart failure | 1 (7.1) |
| Hypertension | 6 (42.9) |
| Peripheral arterial disease | 1 (7.1) |
| Aortic plaques | 1 (7.1) |
| No secondary diagnosis | 5 (35.8) |
| Civil status, n (%) | |
| Single | 1 (7.1) |
| Married or living with a partner | 10 (71.4) |
| Widow or widower | 3 (21.4) |
| Education, n (%) | |
| Unskilled worker | 3 (21.4) |
| Skilled worker | 7 (50) |
| Master’s degree | 4 (28.6) |
| Work status, n (%) | |
| Works under 20 hours/week | 1 (7.1) |
| Works 20 - 36 hours/week | 1 (7.1) |
| Works full time 37 hours/week | 2 (14.3) |
| Retired | 10 (71.4) |

[a]FP-AF: Future Patient—Telerehabilitation of Patients with Atrial Fibrillation.

[b]$CHA_2DS_2VASc$ score: used to assess the risk of stroke in individuals with atrial fibrillation (C: congestive heart failues, H: hypertension, A: age 75 or above (2 points), D: diabetes mellitus, S: stroke/TIA/thromboemolism (2 points), V: vascular disease, A: age 65-74 years, Sc: sex category female).

## Findings

In total, 7 overall themes were generated: "Devices have many functions," "HeartPortal as a digital toolbox," "Coping strategies living with AF," "Measured values are useful," "Community of practice," and "FP-AF program creates a sense of security." Each theme generated several subthemes, which include a representative patient quote identified with an ID number. Table 1 gives an overview of the overall themes and subthemes generated from the interviews.

### Devices Have Many Functions for the Patients

The theme "devices have many functions" explores the patients' experiences with using the various monitoring instruments, which include a blood pressure monitor, weight scale, activity tracker, ECG monitor, and sleep sensor. Overall, patients stated that the devices were user-friendly and easy to set up:

> *I found using the technologies easy; it did not bother me at all. It was easy to use, and did not require great knowledge.* [ID580]

The device mentioned most frequently by the patients was the ECG monitor, as the patients found it to create a sense of security:

> *I missed it when they took it away from me. Being able to measure my heart rhythm was very reassuring.* [ID580]

> *I can imagine that most people, including me, find the ECG important. It was the device I was most pleased about.* [ID666]

The patients also found that the devices motivated them to exercise more. This was particularly true of the step counter:

> *I keep track of my steps during the day. It has become a goal to walk 10,000 steps every day.* [ID726]

These statements indicate that the devices were well-received by patients for their ease of use and functionality. The ECG monitor, in particular, provided reassurance and was valued among the patients. In addition, the activity tracker motivated patients to increase their physical activity, suggesting that these devices can positively impact health behaviors and enhance patient engagement.

### The HeartPortal as a Digital Toolbox

The theme "HeartPortal as a digital toolbox" explores the patients' experiences using the HeartPortal. Several patients found that the HeartPortal functioned like a toolbox for navigating their AF:

> *The HeartPortal has functioned like a toolbox, I used it in connection with the preliminary consultation I had before my ablation, where I watched the videos on the portal.* [ID789]

> *It's nice to have everything gathered in one place. I have been in dialogue with the project nurse a few times, and I used it to keep track of my measurements.* [ID684]

In addition to serving as a toolbox for managing AF, patients found the HeartPortal to be a useful communication platform for communicating with health care professionals. Communication through the HeartPortal was viewed as more time-efficient, allowing patients to ask questions directly without needing to communicate through multiple intermediaries when an immediate answer was not required:

> *I have been in contact with the project nurse through the HeartPortal… We don't have to call and steal each other's time, so I think it's great for those questions where you don't need an answer right away.* [ID674]

> *The platform functions like a direct line where you can skip your doctor and a secretary at the hospital. With medicine and such, I believe that it has helped to make me calmer.* [ID623]

Nevertheless, some patients did not find that they had used the HeartPortal frequently. Some stated that they had not experienced AF during the project and therefore did not feel a significant need for the platform:

> *I have not used the HeartPortal so much during the project. I have written a couple of messages to the project nurse, but I have not had AF during this period, so I have not really felt the need for it.* [ID666]

Others mentioned that they were not accustomed to electronic data processing and therefore had not used the HeartPortal:

> *I have not really used the HeartPortal. I am not used to EDP. I could have done more, but I may not have had the biggest need.* [ID674]

The patients' use of the HeartPortal indicates that the platform is a valuable tool for many patients, as it helps them manage their AF and facilitates easier communication with health care professionals. The communication part of the platform is especially appreciated for non-urgent communication with health care professionals. However, the usage of the HeartPortal varies among the patients, with some patients not feeling the need for it during the period of the study and others being less comfortable using digital communications due to a lack of digital literacy. This suggests that even though the HeartPortal is beneficial for patients, its use may be enhanced by addressing the needs of the individual patient regarding their confidence in the use of digital tools.

### Coping Strategies Living With AF

The theme "Coping strategies living with AF" explores the FP-AF program's effect on patients' coping strategies. Nearly all the interviewed patients reported gaining greater knowledge of how to manage their AF symptoms. One patient stated that the program created a curiosity about her AF:

> *I have learned a lot, and I have gained a greater knowledge of my illness since using the technologies… it has created a curiosity in relation to my illness, in which I've obtained a better understanding.* [ID674]

In addition, some patients stated that their participation in the FP-AF program made them feel more empowered to manage their AF:

> *I feel like I'm capable of handling my AF. Participating in the project has given me a sense of calm.* [ID580]

These findings indicate that the FP-AF program has had a positive impact on patients' abilities to cope with their AF, as it has provided them with increased knowledge of their symptoms and created a sense of security, if not empowerment.

### Measured Values Are Useful

The theme "Measured values are useful" explores patients' use and attitudes toward the data overview provided by their measured values. Patients felt that the measured values created a clear overview of their condition, with one patient noting that this overview even helped her lose weight during the program:

> *I looked a lot at the graphs of my blood pressure and weight, and so on. I actually lost a lot of weight during that time.* [ID580]

Furthermore, the patients also found that having an overview of their data created a sense of security:

> *What I'm doing now, measuring all these values, helps me… I immediately feel better, it has created a sense of security.* [ID684]

> *I couldn't sleep before I got the ECG measure. It created a sense of security being able to see how everything was going.* [ID580]

Of the 14 patients, 2 expressed concern regarding measuring their data. They did not want to become anxious about measuring their values. One patient stated that measuring your values could make you sick:

> *I would not be measuring my values if not for this project. I do not think you should seek out illness, it can make you sick.* [ID686]

These perspectives indicate that while many patients find value in the data overview provided by their measured values, and that this may lead to positive outcomes such as weight loss and increased security, there were also concerns about the potential for increased anxiety and over-monitoring.

### Community of Practice

The theme "Community of practice" explores patients' experiences in the FP-AF program, with a focus on the educational benefits of engaging with peers who also have AF. Many patients found it particularly helpful to learn from and share experiences with others in a similar situation. This community-based learning helped patients gain a greater knowledge of AF and how it affects individuals differently.

> *The dialogue with other patients is a big part of the education at the healthcare centres. You learn a lot by talking to others about how they're feeling and how it affects them.* [ID742]

> *I got so much smarter from being with other people with the same illness. I feel it in my way, but others feel it in a completely different way, something you don't know when you're all alone.* [ID766]

For patients with AF who were already well-informed and not severely impacted by their AF daily, the program's educational aspects were less useful. Some felt they did not gain new knowledge, as they were already comfortable managing their AF:

> *The project has not contributed anything new for me, as I am very oriented about it. I have had AF for many years and acquired a lot of knowledge.* [ID703]

Another patient stated that she felt secure in managing her AF from the beginning and thus did not see the need for the educational program:

> *I felt safe in my illness right away and didn't feel like reading any more about it.* [ID623]

For this patient, participating in the program would only lead to her focusing on issues that were not at all problematic:

> *You can end up focusing on things that might not actually be an issue.*
>
> *[ID623]*

These findings indicate that the community of practice proved beneficial for most patients in enhancing their understanding of AF. Those patients with AF who were more confident in managing their disease may have felt that the program was unnecessary for them compared to others who are less familiar with their condition. Furthermore, there is a concern that participation in the program could lead to unnecessary focus on potential issues of little relevance to the patient. The differences among patients with AF indicate a need for tailored interventions that can incorporate the varying levels of their pre-existing knowledge and experience.

### FP-AF Program Creates a Sense of Security

The theme "FP-AF program creates a sense of security" explores the overall experiences of patients participating in the program. Most patients found the education at the health care center beneficial:

> *I am very satisfied and feel that I have received everything and more during the 4 days of instruction. I don't think there's anything missing or lacking in the course. Not in relation to the teaching'.* [ID789]

Patients also reported that the lectures at the health care center significantly contributed to their sense of security:

> *The sense of security that it has given me is worth its weight in gold.* [ID666]

A few patients noted that the project had given them a sense of calm regarding their AF:

> *At the beginning of my AF, I was in a dark place, and I was scared to walk alone. Today I can walk alone, and I walk 10 kilometres every second day. I am well now… being in the project gave me peace.* [ID580]

Some patients also stated that home monitoring was beneficial because it reduced the need for hospital checkups:

> *You can be monitored at home and don't need to go to the hospital for check-ups so often. This is especially good for elderly people, who may not have a convenient way to transport themselves.* [ID635]

Not all the patients saw the need for the education at the health care center as necessary, however.:

> *I don't take part in the patient education in the healthcare centre. For me, I don't find it necessary to talk about my illness once a week. I've come to terms with it being the way it is.* [ID726]

These perspectives indicate that for most patients, the FP-AF program provides benefits in terms of education, security, and community support. The lectures at the health care center and interactions with other patients help the patients create a sense of security and understanding about their condition. However, the program's value varies among patients, with some not feeling the need for weekly discussions or community engagement. Home monitoring is also found beneficial, particularly for those who find it difficult to travel to the hospital.

## Discussion

### Principal Findings

This study has explored the experiences of patients with AF in participating in the FP-AF program intervention arm. For most patients, the FP-AF program created an enhanced sense of security and empowerment and improved their knowledge of AF. The lectures at the health care center added to their knowledge and sense of security, while the community of practice with peers increased the patients' understanding of the individuality of AF. Patients found the technology user-friendly, and the HeartPortal's data overview further increased their sense of security and motivated them to additional exercise. However, the interviews also revealed that some patients felt that they did not benefit from the education at the health care centers.

These findings are consistent with existing literature that highlights the importance of a community of practice and interaction with health care professionals in the rehabilitation process [27-30]. Kenny et al [27] highlighted the important role of a support network, particularly from staff, and noted patients' dissatisfaction with limited interaction opportunities with peers. Similarly, Anttila et al [28] emphasized that connecting with others is a crucial component of the rehabilitation experience. Furthermore, Lunde et al [29] found that a supportive individual behind a smartphone app is vital for promoting healthy behaviors following rehabilitation.

The study also revealed that the devices used in the FP-AF program were well-received and appreciated for their user-friendliness and functionality. The devices provided patients with a sense of security and motivated them to exercise, indicating a positive impact on health behaviors and patient engagement. These results align with Kenny et al's [27] results, which showed that self-monitoring tools such as heart rate monitors, blood pressure monitors, and activity trackers improved patients' insights into their physical condition and allowed them to track their progress, thereby enhancing their psychological well-being. This is further supported by Olofsson et al [31], who found that self-monitoring enhanced patients' understanding of their symptoms and contributed to a higher level of autonomy [31]. In addition, patients reported that the FP-AF program made them feel empowered to deal with their AF through increased knowledge of their illness, a finding that aligns with Kenny et al's study, which found that CTR empowered patients and resulted in greater knowledge and involvement in their recovery process [27]. Similarly, Su et al [32] found that CTR enhances patient knowledge and empowerment by equipping them with skills to modify their behavior and address everyday challenges.

The HeartPortal was identified as a valuable tool for many patients, serving as a tool for the management of AF and as a platform for easier communication with health care professionals. This finding supports Kenny et al's [27] observation that patients value the ability to contact health care professionals as needed, highlighting the importance of both digital tools and in-person support.

The FP-AF rehabilitation program was based on SDT, where motivation is seen as a key component in any successful rehabilitation [25]. The findings in this qualitative sub-study highlight how the FP-AF rehabilitation program appears to fulfill the needs for autonomy, competence, and relatedness [25]. The program allowed patients to take control of their health, provided patient education and a community of practice, and gave patients useful tools and knowledge to manage their AF. However, a few of the interviewed patients expressed concerns about the self-monitoring aspect of the program, noting that it could lead to excessive focus on issues that may not be relevant, potentially increasing symptoms of anxiety. The relationship between self-monitoring and anxiety is complex, with heterogeneous findings across studies. For instance, Rosman et al [33] found that a self-management program incorporating self-monitoring reduced symptoms of anxiety among patients with AF. Research in other populations, however, such as those with diabetes and heart failure, found no effect of self-monitoring on psychological factors [34,35].

In summary, the study highlights the need for a balanced approach in CR programs that combines digital tools and home monitoring with in-person support and community interactions. A comprehensive rehabilitation program for patients with AF should effectively combine these digital and interpersonal elements to enhance the overall patient experience and outcomes.

The present study also revealed that some patients did not benefit from the FP-AF program. This aligns with Vonk et al's [36] findings that some patients did not participate in CR because they did not feel the need for additional supervision or see benefit in the program's trajectory. Furthermore, Vonk et al [36] found that some participants were reluctant to participate in the group meetings connected to CR, as they did not want to hear about other people's problems. In the present study, some patients were less comfortable with the use of these digital technologies due to limited technical literacy, which may have affected their ability to participate in the FP-AF program. Research suggests that the degree of digital literacy may influence the likelihood of patients using telehealth technologies [37-39]. Moreover, factors such as age, socioeconomic status, race, and eHealth literacy have been found to influence patients' engagement with these technologies [38,40]. Hesitancy to use

telehealth technologies may be rooted in several factors, including difficulties operating the technology, lack of access, poor quality of telehealth appointments, and a preference for in-person care [41]. These insights suggest that while the FP-AF program offers valuable benefits for many patients, its effectiveness may vary based on the individual patient's pre-existing knowledge and experience in managing their condition. Future telerehabilitation for patients with AF needs to tailor interventions to better address the needs of the patients at different stages of managing their condition.

### Limitations

The interviewed participants were recruited with the criteria of having participated in the patient education at the health care centers, which may introduce selection bias. We evaluated patients over a short period of time, and since AF is a chronic condition, a longer follow-up period could have been beneficial for exploring the long-term effects of telerehabilitation. Furthermore, it should be noted that the majority of the interviewed patients had high levels of education, which may have impacted the results, as this patient group may have been more comfortable with the digital toolbox and able to use it more fully. This study is conducted in a Danish context, so the findings might not be applicable in all countries globally. Another limitation is that the interviewers were affiliated with the FP-AF study, which may have introduced response bias, as participants might have responded in ways they thought were more desirable.

### Conclusions

Participation in the FP-AF program enhanced patients' sense of security, empowerment, and knowledge about AF. This improvement was due largely to a combination of patients' use of the Heart Portal and the educational sessions at health care centers. Telerehabilitation for patients with AF may be a useful way of researching this group of patients with a focus on rehabilitation. Telerehabilitation for patients with AF may be an effective means of offering rehabilitation to this group in the future.

## Acknowledgments

We express our gratitude to The Danish Heart Association, Snedkermester Sophus Jacobsen og Hustru Astrid Jacobsens Fond, Aage og Johanne Louis-Hansens Fond, MedCom, Viborg Municipality, Region Midtjylland's Forskningsinnovationspulje, and Rosa og Asta Jensens Fond for providing financial support. We also wish to thank the patients participating in the FP-AF study for participating in the qualitative interviews for this study. Special thanks to the project nurses in the FP-AF study: Helle Mark Mogensen and Hanne Kastbjerg.

## Conflicts of Interest

None declared.

## References

1. Linz D, Gawalko M, Betz K, et al. Atrial fibrillation: epidemiology, screening and digital health. Lancet Reg Health Eur 2024 Feb;37:100786. [doi: 10.1016/j.lanepe.2023.100786] [Medline: 38362546]
2. Vinter N, Cordsen P, Johnsen SP, et al. Temporal trends in lifetime risks of atrial fibrillation and its complications between 2000 and 2022: Danish, nationwide, population based cohort study. BMJ 2024 Apr 17;385:e077209. [doi: 10.1136/bmj-2023-077209] [Medline: 38631726]
3. Staerk L, Wang B, Preis SR, et al. Lifetime risk of atrial fibrillation according to optimal, borderline, or elevated levels of risk factors: cohort study based on longitudinal data from the Framingham Heart Study. Br Med J 2018;BMJ 361:k1453. [doi: 10.1136/bmj.k1453]
4. Elliott AD, Middeldorp ME, Van Gelder IC, Albert CM, Sanders P. Epidemiology and modifiable risk factors for atrial fibrillation. Nat Rev Cardiol 2023 Jun;20(6):404-417. [doi: 10.1038/s41569-022-00820-8] [Medline: 36600003]
5. Van Gelder IC, Rienstra M, Bunting KV, et al. 2024 ESC Guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2024 Sep 29;45(36):3314-3414. [doi: 10.1093/eurheartj/ehae176] [Medline: 39210723]
6. Buja A, Rebba V, Montecchio L, et al. The cost of atrial fibrillation: a systematic review. Value Health 2024 Apr;27(4):527-541. [doi: 10.1016/j.jval.2023.12.015] [Medline: 38296049]
7. Stridsman M, Strömberg A, Hendriks J, Walfridsson U. Patients' experiences of living with atrial fibrillation: a mixed methods study. Cardiol Res Pract 2019;2019:6590358. [doi: 10.1155/2019/6590358] [Medline: 31885904]
8. Bravo P, Edwards A, Barr PJ, et al. Conceptualising patient empowerment: a mixed methods study. BMC Health Serv Res 2015 Jul 1;15:252. [doi: 10.1186/s12913-015-0907-z] [Medline: 26126998]
9. World Health Organization. Health promotion glossary. 1998 WHO/HPR/HEP/98.1.
10. Aujoulat I, d'Hoore W, Deccache A. Patient empowerment in theory and practice: polysemy or cacophony? Patient Educ Couns 2007 Apr;66(1):13-20. [doi: 10.1016/j.pec.2006.09.008] [Medline: 17084059]
11. Werbrouck A, Swinnen E, Kerckhofs E, Buyl R, Beckwée D, De Wit L. How to empower patients? A systematic review and meta-analysis. Transl Behav Med 2018 Sep 8;8(5):660-674. [doi: 10.1093/tbm/iby064] [Medline: 29982675]

12. Bellmann B, Lin T, Greissinger K, Rottner L, Rillig A, Zimmerling S. The beneficial effects of cardiac rehabilitation. Cardiol Ther 2020 Jun;9(1):35-44. [doi: 10.1007/s40119-020-00164-9] [Medline: 31997145]
13. Buckley BJ, Long L, Risom SS, et al. Exercise-based cardiac rehabilitation for adults with atrial fibrillation. Cochrane Database Syst Rev 2024 Sep 17;9(9):CD011197. [doi: 10.1002/14651858.CD011197.pub3] [Medline: 39287086]
14. Palm P, Qvist I, Rasmussen TB, Christensen SW, Håkonsen SJ, Risom SS. Educational interventions to improve outcomes in patients with atrial fibrillation-a systematic review. Int J Clin Pract 2020 Nov;74(11):32726511. [doi: 10.1111/ijcp.13629] [Medline: 32726511]
15. Ruano-Ravina A, Pena-Gil C, Abu-Assi E, et al. Participation and adherence to cardiac rehabilitation programs. A systematic review. Int J Cardiol 2016 Nov;223:436-443. [doi: 10.1016/j.ijcard.2016.08.120]
16. Kotseva K, Wood D, De Bacquer D. Determinants of participation and risk factor control according to attendance in cardiac rehabilitation programmes in coronary patients in Europe: EUROASPIRE IV survey. Eur J Prev Cardiolog 2018 Aug;25(12):1242-1251. [doi: 10.1177/2047487318781359]
17. McDonagh ST, Dalal H, Moore S, et al. Home-based versus centre-based cardiac rehabilitation. Cochrane Database Syst Rev 2023 Oct 27;10(10):CD007130. [doi: 10.1002/14651858.CD007130.pub5] [Medline: 37888805]
18. Owen O, O'Carroll V. The effectiveness of cardiac telerehabilitation in comparison to centre-based cardiac rehabilitation programmes: a literature review. J Telemed Telecare 2024 May;30(4):631-646. [doi: 10.1177/1357633X221085865] [Medline: 35369770]
19. Cai C, Bao Z, Wu N, et al. A novel model of home-based, patient-tailored and mobile application-guided cardiac telerehabilitation in patients with atrial fibrillation: a randomised controlled trial. Clin Rehabil 2022 Jan;36(1):40-50. [doi: 10.1177/02692155211032372] [Medline: 34266323]
20. Pagliari C, Isernia S, Rapisarda L, et al. Different models of cardiac telerehabilitation for people with coronary artery disease: features and effectiveness: a systematic review and meta-analysis. J Clin Med 2024 Jun 10;13(12):38929925. [doi: 10.3390/jcm13123396] [Medline: 38929925]
21. Takahashi EA, Schwamm LH, Adeoye OM, et al. An overview of telehealth in the management of cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2022 Dec 20;146(25):e558-e568. [doi: 10.1161/CIR.0000000000001107] [Medline: 36373541]
22. Dinesen B, Dam Gade J, Skov Schacksen C, et al. The Danish Future Patient Telerehabilitation Program for patients with atrial fibrillation: design and pilot study in collaboration with patients and their spouses. JMIR Cardio 2021 Jul 19;5(2):e27321. [doi: 10.2196/27321] [Medline: 34279239]
23. Dinesen B, Albertsen AE, Joensen EDR, et al. Future patient-telerehabilitation of patients with atrial fibrillation: protocol for a multicenter, mixed methods, randomized controlled trial. JMIR Res Protoc 2025 Feb 18;14:e64259. [doi: 10.2196/64259] [Medline: 39965197]
24. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
25. Ryan RM, Patrick H, Deci EL, Williams GC. Facilitating health behaviour change and its maintenance: interventions based on self-determination theory. Eur Health Psychol 2008;10:2-5 [FREE Full text]
26. Kvale S, Brinkmann S. InterViews: Learning the Craft of Qualitative Research Interviewing, 3rd edition: Sage Publications; 2015. URL: https://uk.sagepub.com/en-gb/eur/interviews/book239402 [accessed 2025-08-08]
27. Kenny E, Byrne M, McEvoy JW, Connolly S, McSharry J. Exploring patient experiences of participating in digital cardiac rehabilitation: a qualitative study. Br J Health Psychol 2024 Feb;29(1):149-164. [doi: 10.1111/bjhp.12692] [Medline: 37722874]
28. Anttila MR, Söderlund A, Sjögren T. Patients' experiences of the complex trust-building process within digital cardiac rehabilitation. PLoS ONE 2021;16(3):e0247982. [doi: 10.1371/journal.pone.0247982]
29. Lunde P, Bye A, Bruusgaard KA, Hellem E, Nilsson BB. Patients' experiences of using a smartphone app after cardiac rehabilitation: qualitative study. JMIR Hum Factors 2022 Mar 23;9(1):e34294. [doi: 10.2196/34294] [Medline: 35319476]
30. Nilsson U, Öberg B, Bäck M. Patients' perceptions of exercise-based cardiac telerehabilitation after a myocardial infarction—a qualitative study. Int J Environ Res Public Health 2023 Apr 6;20(7):37048034. [doi: 10.3390/ijerph20075420] [Medline: 37048034]
31. Olofsson S, Josephsson H, Lundvall M, Lundgren P, Wireklint Sundström B. Patient participation in self-monitoring regarding healthcare of heart failure: an integrated systematic review. BMC Prim Care 2025 Mar 1;26(1):60. [doi: 10.1186/s12875-025-02757-6] [Medline: 40025422]
32. Su JJ, Paguio J, Baratedi WM, Abu-Odah H, Batalik L. Experience of coronary heart disease patients with a nurse-led eHealth cardiac rehabilitation: qualitative process evaluation of a randomized controlled trial. Heart Lung 2023;57:214-221. [doi: 10.1016/j.hrtlng.2022.10.005] [Medline: 36265371]
33. Rosman L, Armbruster T, Kyazimzade S, et al. Effect of a virtual self-management intervention for atrial fibrillation during the outbreak of COVID-19. Pacing Clin Electrophysiol 2021 Mar;44(3):451-461. [doi: 10.1111/pace.14188] [Medline: 33565642]

34. Pekmezaris R, Nouryan CN, Schwartz R, et al. A randomized controlled trial comparing telehealth self-management to standard outpatient management in underserved black and Hispanic patients living with heart failure. Telemed J E Health 2019 Oct;25(10):917-925. [doi: 10.1089/tmj.2018.0219] [Medline: 30418101]
35. Hirani SP, Rixon L, Cartwright M, Beynon M, Newman SP, Team WE. The effect of telehealth on quality of life and psychological outcomes over a 12-month period in a diabetes cohort within the whole systems demonstrator cluster randomized trial. JMIR Diabetes 2017 Sep 1;2(2):e18. [doi: 10.2196/diabetes.7128] [Medline: 30291060]
36. Vonk T, Nuijten MAH, Maessen MFH, et al. Identifying reasons for nonattendance and noncompletion of cardiac rehabilitation. J Cardiopulm Rehabil Prev 2021;41(3):153-158. [doi: 10.1097/HCR.0000000000000580]
37. Chung GS, Ellimoottil CS, McCullough JS. The role of social support in telehealth utilization among older adults in the united states during the COVID-19 pandemic. Telemed Rep 2021;2(1):273-276. [doi: 10.1089/tmr.2021.0025] [Medline: 35720763]
38. McAlearney AS, Sieck CJ, Gregory ME, et al. Examining patients' capacity to use patient portals: insights for telehealth. Med Care 2021 Dec 1;59(12):1067-1074. [doi: 10.1097/MLR.0000000000001639] [Medline: 34593709]
39. Le TV, Galperin H, Traube D. The impact of digital competence on telehealth utilization. Health Policy Technol 2023 Mar;12(1):100724. [doi: 10.1016/j.hlpt.2023.100724]
40. Ezinne NE, Koomson I, Anyasodor AE, et al. Technological and socio-demographic factors influencing telemedicine literacy in Trinidad and Tobago: a cross-sectional study using multidimensional approach. Health Sci Rep 2025 Feb;8(2):e70420. [doi: 10.1002/hsr2.70420] [Medline: 39980826]
41. Kemp M, Rising KL, Laynor G, et al. Barriers to telehealth uptake and use: a scoping review. JAMIA Open 2025 Apr;8(2):ooaf019. [doi: 10.1093/jamiaopen/ooaf019] [Medline: 40109500]


## Abbreviations

**AF:** atrial fibrillation
**CR:** cardiac rehabilitation
**CTR:** cardiac telerehabilitation
**ECG:** electrocardiogram
**FP-AF:** Future Patient—Telerehabilitation of Patients with Atrial Fibrillation
**QoL:** quality of life
**SDT:** self-determination theory


*Edited by KC Wong; submitted 11.11.24; peer-reviewed by KS Menezes, S Risom; revised version received 03.05.25; accepted 15.05.25; published 19.08.25.*

*Please cite as:*
*Joensen EDR, Albertsen AE, Spindler H, Jensen KM, Frost L, Dittmann L, Gunasegaram M, Johnsen SP, Jochumsen MR, Svenstrup D, Dinesen B*
*Patient Perspectives on the "Future Patient" Telerehabilitation Program for Atrial Fibrillation: Qualitative Study*
*JMIR Cardio 2025;9:e68663*
*URL: https://cardio.jmir.org/2025/1/e68663*
*doi:10.2196/68663*

# mHealth Support in Cardiac Care Pathways for Patient Self-Management During Transitions From Hospital to Rehabilitation: Exploratory Field Study

Isabel Höppchen[1]; Stefan Tino Kulnik[1]; Alexander Meschtscherjakov[2]; Josef Niebauer[1,3,4]; Bernhard Reich[3,4]; Jan David Smeddinck[1]; Daniela Wurhofer[1]

[1]Ludwig Boltzmann Institute for Digital Health and Prevention, Lindhofstr. 22, Salzburg, Austria

[2]University of Salzburg, Salzburg, Austria

[3]University Institute of Sports Medicine, Prevention and Rehabilitation, Paracelsus Medical University, Salzburg, Austria

[4]REHA Zentrum Salzburg, Salzburg, Austria

**Corresponding Author:**
Isabel Höppchen
Ludwig Boltzmann Institute for Digital Health and Prevention, Lindhofstr. 22, Salzburg, Austria

## Abstract

**Background:** Cardiac rehabilitation (CR) is essential for recovery from cardiovascular disease. However, patients often encounter challenges in navigating the transition from acute hospital care to CR. Mobile health (mHealth) technologies may support this critical phase; however, evidence regarding their clinical practice remains limited. The HERO app (developed by REDOX GmbH) was developed to address the needs of patients with cardiovascular disease for orientation, emotional support, and motivation during this transition.

**Objective:** This study aims (1) to explore how mHealth technologies tailored for patients with cardiovascular disease can support their needs regarding orientation, emotional balance, and motivation during the transition from the acute hospital to CR and (2) to evaluate the user experience and acceptance of the HERO app as targeted pathway support.

**Methods:** A mixed methods study was conducted with patients with cardiovascular disease using study diaries, questionnaires, and semistructured interviews. Participants were purposively recruited in acute hospitals and rehabilitation settings. Quantitative data were analyzed descriptively, and qualitative data were analyzed using content analysis after Mayring.

**Results:** Eight participants used the app for an average of 14 (range 4-23) days. The app was perceived as a helpful short-term resource. It supported patients in understanding their condition, planning for CR, and regaining motivation. Participants highlighted the value of combining objective information with peer experiences. Suggestions for improvement included more personalized self-management guidance and a precise onboarding process to increase accessibility and usability.

**Conclusions:** Based on the findings, we propose 4 pillars of mHealth support for cardiac care transitions, including timely access, actionable guidance, peer support, and short-term usability. These pillars could inform the design of patient-centered mHealth tools for care transitions.

***(JMIR Cardio 2025;9:e76089)*** doi:10.2196/76089

**KEYWORDS**

cardiac rehabilitation; mHealth; digital health; empowerment; patient transitions; telemedicine; transitional care; usability; mobile health

## Introduction

Cardiac rehabilitation (CR) is highly recommended by international guidelines as a class 1 indication for many patients with cardiovascular disease, for example, after an acute myocardial infarction [1]. However, participation rates are low, with approximately only 34% of eligible patients globally taking up CR [2]. Missing out on rehabilitation can have profound implications on patients' long-term morbidity and mortality [3]. Early uptake of CR after hospitalization is beneficial to mitigate potential uncertainty about the cardiac condition, for example, by learning coping strategies [4,5]. Nevertheless, transitions from the acute hospital to rehabilitation often include multiple steps and can be complex and challenging for patients [6,7].

In a previous study, we found that after discharge from the acute hospital, patients experience primarily three needs: (1) a "need for emotional balance" to reflect on the cardiac event and the time in the hospital; (2) a "need for orientation" to get to know and consider follow-up care options, such as CR; and (3) a "need for motivation" for a healthy lifestyle, including the participation

in CR [8]. Adequate and timely information can help address these needs, but clinical practice to date fails to address these needs and implement effective solutions [9,10].

Information provision should ideally commence early after an acute cardiac event to increase patients' confidence in managing their condition and improve the likelihood of CR uptake [10]. However, access to reliable information sources is often limited to the patients' hospital stay, as health care professionals are their primary information sources [11,12]. Previous research indicated a mismatch between the amount of information health care professionals provide and patients' ability to capture it [13]. Information provision during hospitalization is also often provided verbally and low-technology methods, constraining the possibility of rereading and processing beyond patients' hospital stay [13].

After discharge, patients often struggle with self-assessing their physical capacity and adopting healthy behaviors, such as dietary changes, exercising, and smoking cessation [14]. During this critical step in the patient pathway, patients need to know how to self-manage their condition. Peer support, goal setting, and access to health services can be adequate resources for patient support [13]. Mobile health (mHealth) technologies have the potential to provide access to reliable information beyond patients' acute hospital stay and, therefore, widen the time window for processing information [13]. They could offer tailored information according to the rehabilitation progress, adapted to patients' health literacy and individual care pathways across health care sectors [8,11,15].

Several studies found evidence that mHealth could increase patients' knowledge and, thereby, empower them to actively participate in their care process and decision-making [16-19]. Vardoulakis et al [19] provide an example of how a smartphone app can affect patients' knowledge by providing layman-friendly health information. The interaction with the app increased patients' knowledge about their care plan and allowed them to actively participate during daily ward rounds, for example, by asking specific questions [19]. Also, the patients appreciated that the app removed dependence on health care professionals to provide information. Being aware of the possibility to reread details had a calming effect on them [19].

Despite the potential for effective mHealth support, there is a lack of evidence about digital support for intersectoral, multistep patient pathways. Related studies often focus exclusively on specific health care touchpoints, such as outpatient care or long-term lifestyle adjustments [20-22]. Less is known about the targeted support needed during phases when patients find themselves between 2 health care touchpoints. In light of this, we developed the "Dein Weg zur Reha" ("Your pathway to rehabilitation") smartphone app (HERO app) to support the evolving needs of patients with cardiovascular disease along their care pathway [8].

This exploratory field study investigated how an mHealth technology, such as the HERO app, could address the informational and emotional needs of patients during critical transitions from the acute hospital to CR. The objectives of this study were (1) to explore how mHealth technologies tailored for patients with cardiovascular disease can support their needs regarding orientation, emotional balance, and motivation during the transition from the acute hospital to CR, and (2) to evaluate the user experience and acceptance of the HERO app.

## Methods

### Overview

This exploratory field study investigated how the HERO app could support patients with cardiovascular disease during their transition from the acute hospital to CR. To capture user experiences from a naturalistic use setting and to assess the app's sociotechnical context [23,24], we integrated the app into the patient pathway after hospitalization.

### The HERO App for Tailored Patient Support

This study builds on previous research, including a contextual inquiry [6] and an in-depth understanding of the needs of patients with cardiovascular disease along care pathways [8]. The results informed the development of the HERO app, which aims to address patients' needs along their care pathway by providing necessary and reliable information. Figure 1 illustrates the HERO app interface and its application context within the patient pathway, mapped along the domains of the NASSS (Nonadoption, Abandonment, Scale-Up, Spread and Sustainability) framework, including technology, users, organization, and wider system. Multimedia Appendix 1 presents a video walkthrough of the app.

**Figure 1.** The HERO app and its use context. mHealth: mobile health.


In the following, we describe the HERO app according to the domains of the NASSS framework, developed by Greenhalgh et al [25]. The NASSS could be used to understand and evaluate the challenges of implementing technologies in health care systems.

The app was developed using React Native and is provided as an Android Package Kit (mobile app). It requires a minimum operating system version of Android 5.0. Users must have an Android smartphone and an active internet connection to access the app and linked content.

The target users of the HERO app are individuals with cardiovascular disease for whom CR is medically indicated and recommended. Informal caregivers and relatives could also benefit from the information provided in the app. In addition, health care professionals may find the app valuable as it offers a time-saving method for delivering information and educational content during daily ward rounds. The app is designed to provide self-explanatory content and intuitive navigation, ensuring that health care professionals do not need to spend time introducing patients to its use.

The HERO app is a standalone mHealth technology and can be used offline and without entering personal data to avoid potential privacy concerns. The decision to design the app without requiring interoperability with other technologies on a patient ward was intentional to be easily downloaded onto patients' smartphones without complex integration into hospital infrastructure.

The home screen presents a visualization of a patient pathway from hospital to CR, including the 3 steps "Hospital," "Home," and "Rehabilitation" [26]. Within these 3 steps, users find evidence-based information about CR and experiential knowledge in videos with testimonials from 2 former CR participants [27,28]. The app offers most content bilingually in German and English to accommodate users with different language preferences within the Austrian health care context. The content changes from step to step to meet the temporal dimension of patients' information needs and processing [8,11]. We also included a "backpack" function, in which users can save preferred information and take notes, taking these with them from one step to the next.

## Recruitment Strategy and Inclusion Criteria

We recruited patients with cardiovascular disease between August 2024 and March 2025. Following a purposive sampling strategy, participants were recruited by 2 gatekeepers in hospitals located in different federal states in Austria. The gatekeepers were personal contacts of the project lead and senior physicians in cardiology wards. The gatekeepers screened patients against the inclusion criteria and provided them with a flyer about the study, contact details of the project lead, and a web link to sign up for the study. Patients could also give their telephone

numbers and agree to be contacted by the project lead for recruitment purposes. After the patients gave written consent to participate in the study, they were provided with a study diary and a link to download the app on their smartphones.

Inclusion criteria were hospitalization due to a recent acute cardiac event, a newly diagnosed cardiovascular disease with an indication for attending or currently attending CR in Austria, using a smartphone with an Android operating system, and willingness to install the HERO app on that phone. Exclusion criteria were using a smartphone with an iOS operating system, age less than 18 years, inability to give informed consent, or inability to participate due to the health condition. The recruitment proved to be exceptionally challenging in the hospital setting. A notable barrier was the limitation that the HERO app was only available for Android. Many eligible patients used iPhones (Apple Inc) and declined the offer to borrow an Android phone during the study period. Further reasons included participation in other studies and limited physical or mental energy to participate in research. In some cases, patients could not be reached again after initially expressing interest. Therefore, we decided to broaden the inclusion criteria to also include individuals who were currently participating in a CR program in an outpatient rehabilitation center.

### Quantitative and Qualitative Measures

For the app-testing phase, participants received a paper-based study diary. It included onboarding activities to get to know the features of the app, such as “Save a relevant information card in the backpack” or “Go to the ‘Hospital’ and watch a video.” Participants were asked to complete the AttrakDiff questionnaire for user experience [29], the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) questionnaire for user acceptance [30], and a questionnaire about the frequency of app use. In addition, open questions and space for notes were included to encourage reflection on the overall user experience and interaction with the HERO app. After the app-testing phase, we conducted semistructured interviews with the participants to capture their experiences with the app and to explore to what extent the app had an impact on their pathway to CR. For participants who were recruited during their CR program, interview questions were framed to encourage hypothetical reflection on how the HERO app might have supported them during the earlier transition phase. The interview guide can be found in Multimedia Appendix 2. All interviews were conducted by the project lead and were audiotaped. The interview language was German, and illustrative quotes for the “Results” section of this paper were translated into English.

### Dataset and Analysis

The quantitative dataset consisted of the AttrakDiff, UTAUT2, and frequency of use questionnaire. This data were analyzed descriptively using Excel software (Microsoft). The qualitative dataset consisted of participants’ diary entries and 193 minutes of interview recordings (mean 24; range 13-40 minutes). Audio recordings were transcribed verbatim by the project lead and supported by f4x software (Dr. Dresing & Pehl GmbH). To protect participants’ confidentiality, identifying information was removed and data were pseudonymized prior to analysis. The transcripts were analyzed following qualitative content analysis after Mayring [31], a structured approach to systematically categorize qualitative data. To begin, we defined a coding guide with overarching categories (Multimedia Appendix 3). These categories were derived from our research questions, our previous study on patient needs [8], and the perception of user experience provided by Hasselzahl et al [32]. Categories were then applied to the data and further refined inductively into subcategories to achieve a fine-grained analysis. In line with Mayring [31], we defined the “coding unit” [31] as multiple words sharing a common meaning to ensure a consistent coding approach. The “context unit” [31], which provided interpretive background, included the entire interviews, diary entries, and demographic questionnaires. Finally, we considered the “analysis units” [31] to be the diary entries and the questionnaire responses. The paraphrased transcripts were then coded with the category system by the project lead, supported by MAXQDA 2022 software (VERBI Software). The category system and interpretations were discussed within the research team to ensure consistency and reliability.

### Ethical Considerations

The study was reviewed by the research ethics committee of the Ludwig Boltzmann Gesellschaft in Austria and received a favorable opinion (reference 014_2024). All patients gave written consent to participate in the study. After study completion, they received a compensation of €30 (US $35.17).

## Results

### Overview

In the following, we present participant characteristics (Table 1) and overall feedback on the HERO app’s usability, aesthetics, and use context. We provide insights into the five key themes that emerged from the data: (1) Considering and preparing for cardiac rehabilitation, (2) enhancing motivation to return to daily activities, (3) supporting sense-making of the cardiac event, (4) ensuring timely and location-independent access to the app, and (5) improving self-management guidance and addressing information gaps.

**Table .** Participant characteristics and interview settings.

| ID | Sex | Age (years) | Profession | Cardiovascular diagnosis | Year of diagnosis | Participated in CR[a] | Interview setting |
|---|---|---|---|---|---|---|---|
| P01 | Male | 54 | Insurance employee | Myocardial infarction, status post stent implantation | 2024 | Yes (first time) | Face-to-face interview at his home before he started CR |
| P02 | Male | 59 | Company technician | Myocardial infarction, status post stent implantation | 2024 | Yes (first time) | Telephone interview before he started CR |
| P03 | Male | 74 | Artist, retired teacher | Myocardial infarction, status post stent implantation | 2019 | No | Telephone interview after he decided not to participate in CR |
| P04 | Female | 54 | Kindergarten teacher | Acute coronary symptom (NSTEMI)[b] | 2024 | Yes (first time) | Telephone interview during CR |
| P05 | Female | 32 | Teacher | Peripartum cardiomyopathy, functional mitral insufficiency | 2024 | Yes (first time) | Face-to-face interview during CR |
| P06 | Male | 66 | Retired coach for corporate development | Myocardial infarction, status post stent implantation | 2012 | Yes (second time) | Face-to-face interview during CR |
| P07 | Female | 77 | Retired medical technical assistant | Coronary heart disease | 2012 | Yes (fourth time) | Face-to-face interview during CR |
| P08 | Male | 57 | Logistics manager | Coronary heart disease | 2024 | Yes (first time) | Telephone interview during CR |

[a]CR: cardiac rehabilitation.

[b]NSTEMI: non-ST-elevation myocardial infarction

## Participant Characteristics

The study population consisted of 8 participants who were recruited in 3 different federal states in Austria. The mean age was 59.1 (range 32-77) years, and the majority were male (5 self-identifying males and 3 self-identifying females). A total of 4 participants were recruited during hospitalization and 4 during their CR program. While 2 participants were referred by the hospital staff, 4 were referred by their general practitioner or by their internist, and 1 initiated and administered the referral himself. For CR attendants who were hospitalized, the average time between discharge and the start of CR was 69 (range 24-165) days. Overall, 5 participants attended CR for the first time, 2 had attended CR previously, and 1 declined CR participation.

## Overall Feedback on Usability, Aesthetics, and Use Context

The participants used the HERO app for 14 (range 4-23) days on average. Users' UTAUT2 and AttrakDiff ratings are presented in Table 2. The users stated that the app use slightly became a habit (median 4, IQR 1-5).

**Table .** Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) and AttrakDiff ratings from study participants (n=8).

| Questionnaire (score range) and dimension | Median (IQR) |
|---|---|
| UTAUT2 (range 1 to 7, higher score indicating better user acceptance) | |
| Habituation | 4.0 (1-5) |
| Perceived usefulness | 4.5 (2-6) |
| Acquired knowledge | 6.0 (3-7) |
| Intention to use | 2.5 (1-6) |
| Experienced enjoyment | 5.0 (4-6) |
| AttrakDiff (range –3 to 3, higher score indicating better user experience) | |
| Pragmatic quality | 2.0 (–3 to 3) |
| Hedonic quality–stimulation | 0.5 (–3 to 3) |
| Hedonic quality–identity | 1.0 (–1 to 3) |
| Attractiveness | 1 (–1 to 3) |

The low intention to use the app in the future (median 2.5, IQR 1-6) contrasts with generally positive ratings for usefulness and enjoyment (Table 2). This arguably reflects the app's short-term intended use rather than dissatisfaction. Some individuals described the app as a "tool" (P06) used mainly between discharge and the start of CR, after which it could be set aside. For example, participant "P04" indicated that she consulted the app multiple times at home but discontinued use once CR started, as she no longer had the time or a need for it.

Most participants reported having the necessary knowledge to use the app (median 6, IQR 3-6). Only participant "P03" expressed skepticism about digital technologies in general and a lack of familiarity with installation and navigation. The pragmatic quality ratings varied (median 2, IQR –3 to 3), indicating different perceptions of user-friendliness and functionality. Most participants emphasized the app's straightforward design and ease of use, noting that its user guidance was largely self-explanatory. This would make the app "easy to handle even for patients in a hospital bed" (P05). Participant "P06" reported a technical difficulty, as he was unable to open a web link in the app. Regarding the app design, some participants found the mascot, graphics, and color scheme "visually appealing and beautiful" (P08) while others considered that older users may not connect with the design. These opinions were reflected in the attractiveness ratings (median 1, IQR –1 to 3).

## Considering and Preparing for Cardiac Rehabilitation

Both the videos and informational content helped users understand CR as a key component of their recovery from the cardiac event and validated the decision to attend CR or not. The representation of the pathway signposted out the way to CR step by step, offering a sense of security and direction. Participants "P05" and "P08" proposed to even extend the pathway and include tips for navigating the phase following a CR. For most participants, however, the app represented a guide for CR preparation. The note-taking feature enhanced this by nudging them to record relevant details. Participant "P01" described being called by the CR center and told his program would start in a few days. Overwhelmed, he used the app to gather necessary information and write a packing list:

> *The app was really helpful. I got admitted so quickly. […] For me it was just: hurry up and pack! […] So I sat down, looked through the app, and it showed me everything I needed.* [P01]

The link to the "rehabilitation compass," an official website in Austria that informs about rehabilitation options near one's location, supported some participants in selecting the most suitable option. This feature was especially relevant for participant "P03", who decided not to participate in a CR program because the only rehabilitation center near his home did not offer CR services:

> *The compass was very helpful for me because it showed where the CR centers are located. I saw there was a rehabilitation center near me, but it was not an option because it does not have a cardiovascular department.* [P03]

## Enhancing Motivation to Return to Daily Activities

Some participants noted that after a cardiac event, one might either try to ignore it or overanalyze its causes. Therefore, they appreciated that the videos addressed uncertainties and fears after a cardiac intervention, showing how the individuals dealt with those challenges. For example, participant "P04" shared that the videos encouraged her to return to her gym training and to talk to relatives about her cardiac event:

> *The video really got me thinking. That was when it finally sank in what actually happened to me. It was a real 'aha' moment. I had never really understood it before. […] It suddenly hit me that I almost died, and you have to process that. It was awful. […] I had a day or two where I did not want to be around anyone. It felt really strange. That is why it is so important to talk with people, you need to let it all out so you can handle it better.* [P04]

Further, the HERO app motivated some participants to set goals for the phase after hospital discharge and for CR. For example, participant "P06" mentioned that using the app taught him "to

start setting daily goals at home, to make sure I do something, to get moving. That was an interesting realization for me" (P06). The displayed pathway further motivated the participants by showing the way back to "normality" (P01) and fostering a sense of optimism that "everything will be okay, things are looking up" (P04). It also led to a self-reflection, as expressed by participant P05: "It made me realize how much I've already accomplished."

### Supporting Sense-Making of the Cardiac Event

Overall, the video content was perceived as valuable for peer connection, helping participants navigate their rehabilitation journey beyond medical information alone. Some reported that the videos provided a sense of belonging and peer support in moving past the constant worrying about the cardiac event:

> *[The videos] show that you're not alone in what you're going through, and I think that's really, really important* [P07]

> *The videos showed me that other young and fit people had gone through the same thing. That really helped me break out of my cycle of overthinking* [P05]

Some participants also shared that hearing about more severe cardiac diseases made them reconsider the severity of their own experience:

> *Listening to others who went through worse made me think: Okay, my case wasn't that bad after all.* [P04]

While some participants could easily relate to the portrayed individuals, others found them less representative of their situation. In line with this finding, one of the most frequently mentioned ideas was to include more videos featuring different individuals, such as diverse age groups, genders, and cardiac conditions, allowing users to relate more closely to those experiences:

> *The guy basically has my story, he had a heart attack, and probably has cholesterol problems, but he's super skinny! […] I would have liked to see someone who has my vibe [laughs]. Someone overweight. […] It would be cool if, when I open the app, I could see someone who is in my demographic.* [P06]

### Ensuring Timely and Location-Independent Access to the App

Most participants found the HERO app useful in their everyday lives (median 4.5, IQR 2-6) and generally enjoyed using it (median 5, IQR 4-6). All participants emphasized that earlier access to the app would have been even more beneficial; they stated that they would have liked to use it during their hospital stay to receive early information about follow-up care. Some participants had already actively searched for information on the web during their hospitalization or were familiar with the rehabilitation program from previous experiences. As a result, they reported that the app provided little new content for them:

> *I started gathering information while I was still in the hospital […]. I had a heart attack. It came as a big surprise, given that I do a lot of hiking and cycling. I'm not even 60 yet, turning 60 in two weeks. […] So, I spent a lot of time looking into what I could do and what would happen next, even before I had the app. That's why I already knew quite a bit.* [P02]

The participants emphasized that having the app readily available on a smartphone was advantageous, as it allowed for location-independent access at any time. They used the app both at home and on the go. For example, participant "P01" described using the note-taking function while shopping for sportswear in preparation for his CR stay:

> *You are out shopping, you have got your packing list written down, and you open the app like, "Oh, right, I need this, that, and that." It is awesome. […] I do not have to walk around juggling a bunch of handwritten notes anymore, it is all in one place.* [P01]

### Improving Self-Management Guidance and Addressing Information Gaps

Some recommendations in the HERO app were considered too vague, leading to dissatisfaction. For example, the information to ask insurance providers or general practitioners was perceived as not being helpful. Participants feared being stuck in phone queues or that health care professionals would not take their concerns seriously, particularly when they felt uncertain about managing their cardiac condition. Therefore, they requested more concrete information regarding lifestyle and coping strategies for the posthospital phase, with detailed tips on exercise, nutrition, and stress management. Participants suggested a feature for storing medication details, mentioning the possibility of side effects, and health care professionals' contact details. Also, a direct download link for the CR referral form was suggested. Some wanted clearer insights on insurance and cost coverage. Several participants emphasized the importance of providing novel information in the app, offering details not readily available on rehabilitation centers' websites. For example, they suggested including positive messages to reinforce the idea of turning the cardiac event into an opportunity to adopt a healthier lifestyle.

## Discussion

### HERO App as a Short-Term Tool to Bridge Patient Transitions

This exploratory study researched how the HERO app, an mHealth technology designed to guide patients from hospital discharge to CR, can support individuals during a critical care transition. The app was mainly perceived as a short-term tool to bridge the phase after hospital discharge, after which it can be set aside again. The participants emphasized the value of accessing the app early, ideally during their hospital stay, to receive timely information. They also valued the combination of evidence-based information about CR and experiential knowledge provided by peers to self-manage their patient pathways effectively. Patients felt supported in their decision-making about CR participation by receiving objective information about CR, which was presented as one option for follow-up care without pressuring them to participate. This is in line with the understanding of shared decision-making

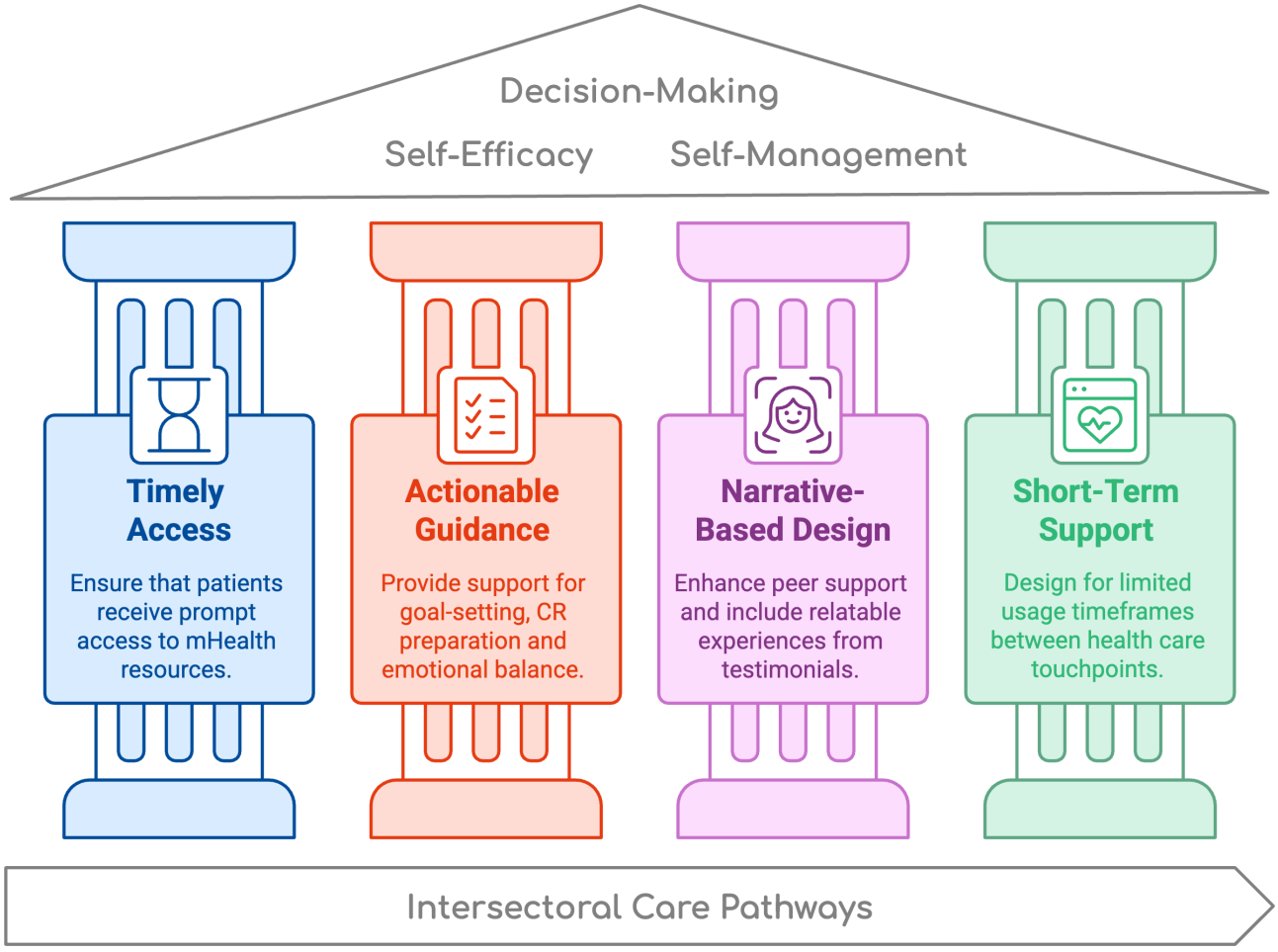
described by Elwyn et al [33], stating that no participation is also a valid option. We also identified areas for improvement, particularly in the personalization and concreteness of self-management support.

## Implications for Integrating mHealth Into Transitional Care Contexts

### *Overview*

Based on our findings, we derived four implications for designing and integrating mHealth technologies for patient support between 2 health care touchpoints: (1) Ensure timely access to mHealth support, (2) provide actionable guidance for self-management, (3) include peer narratives for emotional reassurance, and (4) design for short-term time frames. Figure 2 visualizes the implications as 4 pillars that support patient decision-making and can enhance their self-efficacy and self-management skills. The model could inform the development of patient-centered digital technologies by addressing their need for timely, actionable, and emotionally supportive guidance. Future research could apply this model to design, adapt, or evaluate mHealth interventions across different medical contexts where patients face similar challenges in managing care transitions.

**Figure 2.** Four pillars of mHealth for cardiac care that can support patient decision-making, self-efficacy, and self-management during intersectoral care transitions. CR: cardiac rehabilitation; mHealth: mobile health.




### *Implication 1: Ensure Timely Access to mHealth Support*

Our study highlights the importance of early access to information provided by mHealth, as participants expressed a need for immediate guidance after a cardiac event to support emotional processing, orientation, and the initiation of self-management while being hospitalized. Due to the recruitment process, we were only able to provide the app to patients after their discharge from the acute hospital. For some, this was too late, and they had already gathered information through other sources. To ensure timely access, we argue that mHealth technologies, such as the HERO app, should be integrated into patient pathways and discharge protocols. For example, a combination of automated referral technologies and mHealth support for patients may help to initiate patient pathways within the hospital setting [34]. Targeted implementation strategies based on a holistic context analysis provided by the NASSS framework, for example, could provide support here [25,35].

### *Implication 2: Provide Actionable Guidance for Self-Management*

A key finding of this study was that participants expressed a strong need to actively manage their recovery and navigate their care pathway with a cardiac disease. To achieve this, having all information stored in 1 place was perceived as useful, highlighting the potential of mHealth in reducing patients' burden of searching for and storing information from multiple

sources [36,37]. Providing step-by-step guidance for organizing daily life and tailored information according to the pathway progress can support the navigation of health and administrative challenges. Our study showed that patients prefer concrete, actionable guidance over general recommendations; they request tips on physical activity, nutrition, stress management, and administrative steps for CR access. However, for patient safety, some types of advice, such as information on individual medication plans, should be personalized and provided by qualified health care professionals only. We also argue that, from a practical standpoint, the more information is incorporated into the app, the more it needs to be updated and quality-checked on an ongoing basis.

Nevertheless, the need for such content highlights a readiness to engage in self-management but also a reliance on (digital) tools to do so effectively. Participants emphasized that the app should display information in addition to what they receive from websites and health care professionals, indicating that they trust the app's content and consider the displayed information as reliable. The HERO app offered a foundation for self-management by providing a visual CR pathway, a note-taking feature, and videos from former rehabilitation participants. These elements helped participants reflect on their condition, set goals, and regain a sense of agency and empowerment. According to the health action and process approach [38], a model for understanding health behavior change, forming intentions is based on risk perception and self-efficacy, factors that were activated by participants' engagement with the app. For instance, several participants described the video content as eye-opening, helping them realize the seriousness of their condition and set small goals, such as organizing logistics for CR or integrating physical activity into their daily routines. In line with this, Schneider-Matyka et al [39] found evidence that information influences patients' coping strategies, leading to less avoidance and more active engagement with their recovery.

### *Implication 3: Include Peer Narratives*

The videos from testimonials emerged as a central feature supporting patients' emotional balancing and coping with their cardiac disease. On the one hand, the videos fostered a sense of belonging, helping users feel less isolated after their diagnosis that is confirmed by previous research [40]. Furthermore, participants reported that the videos prompted self-reflection and emotional processing, helping to overcome uncertainties about physical activity and enter the CR program feeling better prepared. Active engagement with experiences helped participants stop overthinking the cardiac event. This is considered a critical step before returning to daily activities, which builds the basis for self-efficacy and long-term behavior changes [38,41].

On the other hand, we found that participants' identification with the testimonials varied, and some participants felt unrepresented due to differences in age, body type, or cardiac history. Sillence et al [28] state that patients could reject information if it does not resonate with their own experiences or health conditions. Based on our findings, we partly agree, but we also want to highlight an additional facet of peer support. We found that contrasting different conditions can be helpful for some patients. The videos encouraged social comparison, allowing our participants to contextualize their condition. To enhance the potential of peer support, diversity in age, gender, and diagnosis should therefore be prioritized to ensure broader identification and inclusiveness [27,42].

### *Implication 4: Design for Short-Term Support*

One interesting pattern we observed was that participants rated the app as useful and enjoyable; however, they expressed low intention to continue using it. The qualitative data help explain this disconnect. Many participants viewed the HERO app as a short-term tool to support the period between hospital discharge and the start of CR. Once CR began, they no longer felt the need to engage with the app. This supports the idea that mHealth tools for care transitions may not need long-term engagement to be effective but instead should focus on timely, targeted support. Unlike long-term mHealth interventions designed to sustain behavior change, mHealth used by patients in transition phases should include tailored content that can be quickly accessed with minimal onboarding and designed for limited usage time frames. Tools, such as the HERO app, aim to activate self-management and promote readiness rather than facilitate continuous tracking or habit formation. This distinction highlights the importance of tailoring mHealth technologies to short-term engagement, for example, by supporting brief patient interactions with tools for quick reference or note-taking [19].

## Limitations and Implications for Future Research

Our decision to close data collection after 8 participants was based on several considerations. First, we were able to recruit a sample, including variation regarding age, gender, years of living with cardiovascular disease, perspective on the patient pathway, and previous CR experience. Following Guest et al [43], who found that thematic saturation often occurs between 6 and 12 interviews in homogenous qualitative studies, we concluded that our 8 interviews were sufficient to capture the main themes. Recurrent themes appeared across the participants, indicating that our research questions had been comprehensively addressed. Nevertheless, participants' experiences were shaped by the specific Austrian health care context, which may affect transferability to other systems. Second, data analysis was conducted in parallel with continued interviewing. This process allowed us to constantly assess the data richness and realize the point of data saturation when additional data would likely lead to redundancy rather than new findings. Third, recruitment in the acute hospital setting proved to be exceptionally challenging due to patient availability. Another significant barrier to participant enrollment was that the app was only available for Android devices, excluding iOS users from participation. Consequently, several interested individuals were unable to participate, which may have introduced a selection bias favoring Android users. Future iterations of the HERO app are planned to include iOS compatibility to ensure broader accessibility. Given these limitations, we assessed data saturation pragmatically, consciously balancing the data richness with recruitment constraints.

Furthermore, our recruitment approach may have resulted in selection bias regarding language skills, formal education levels,

and personal interest in digital tools. Consequently, the perspectives of patient groups typically at risk of being underinformed, such as those with lower education levels or non-German-speaking backgrounds, may not be fully represented.

To ensure app uptake directly from a hospital bed, further refinements of the HERO app should improve accessibility, for example, by offering a more comprehensive onboarding process that explains key app features in greater detail, such as through a video walkthrough. This could improve ease of use and strengthen both initial engagement and perceived usefulness [20]. Due to budgetary and technical constraints, the current HERO app prototype does not include condition-specific content (eg, for patients recovering from cardiac surgery). Future iterations should incorporate more individualized information to reflect different patient experiences and care needs. Research shows that personalized content increases engagement with mHealth tools, and tailoring content to individual needs is a key factor in sustained use [20]. Adapting the HERO app for specific patient groups, such as postsurgery versus patients after a myocardial infarction, could therefore enhance both its relevance and long-term impact.

Following these improvements, a structured evaluation of the effects of the HERO app on pathway adherence, patient activation, and health literacy would be valuable. A randomized controlled trial with extended follow-up periods would be well-suited for this purpose. In addition, targeted implementation strategies could be developed to support the integration of mHealth tools into clinical routines [35]. Combining sociotechnical frameworks and user-centered design principles could be a promising approach [44]. Finally, future studies should examine the role of health care professionals in integrating mHealth into practice and in supporting patients in using such tools effectively [45].

To conclude, we would like to discuss the applicability of our 4-pillar model to other medical fields in which patient transitions between health care settings are relevant, such as orthopedics. While the challenges faced by patients may vary, such as the fear of mortality in patients with cardiovascular disease versus mobility concerns in patients with orthopedic conditions, both groups share insecurities about managing their new health conditions. Scott et al [46] highlighted that patients with orthopedic conditions could experience depression and anxiety, significantly impacting their overall recovery and quality of life. In this regard, the 4 pillars identified in our study—timely intervention, actionable guidance, relatable peer narratives, and short-term support—could also provide valuable support in developing targeted mHealth interventions for orthopedic conditions. However, it is crucial to identify target group-specific needs and contextual factors before applying our findings to other medical fields.

## Conclusion

This exploratory field study investigated how mHealth technologies, such as the HERO app, can support patients with cardiovascular disease in navigating care transitions from the acute hospital to CR. The app helped participants "break out of the cycle of overthinking" (P05) by providing short-term support, enabling self-management, and supporting decision-making and self-efficacy. User experiences suggested that the HERO app could support patient transitions, particularly in considering and preparing for CR, enhancing motivation to return to daily activities, and making sense of the cardiac event. Based on the findings, we propose four pillars of mHealth support for cardiac care transitions: (1) timely access to information, (2) actionable guidance for self-management, (3) peer narratives for emotional reassurance, and (4) design for short-term usability. This model may guide the future design and implementation of mHealth tools across various medical contexts where patients navigate complex care transitions.

## Acknowledgments

We would like to thank Prof Dr Andrea Podolsky, Dr Sebastian Globits, Dr Michael Porodko, and Dr Jakob Nimpf for their exceptional support and valuable contributions to this study. We used the generative artificial intelligence (AI) tool ChatGPT by OpenAI to translate participant quotes from German to English (selectively and after anonymization). The translations were reviewed and revised by the project lead. We used the generative AI tool Napkin AI to draft Figure 2. The visualizations were refined by the project lead.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
HERO app video walkthrough.
[MOV File, 13122 KB - cardio_v9i1e76089_app1.mov ]

Multimedia Appendix 2
Guide for semistructured interviews with study participants.
[DOCX File, 16 KB - cardio_v9i1e76089_app2.docx ]

Multimedia Appendix 3

Category system for data analysis.
[DOCX File, 17 KB - cardio_v9i1e76089_app3.docx ]

## References


1. Virani SS, Newby LK, Arnold SV, et al. 2023 AHA/ACC/ACCP/ASPC/NLA/PCNA guideline for the management of patients with chronic coronary disease: a report of the American Heart Association/American College of Cardiology Joint Committee on clinical practice guidelines. Circulation 2023 Aug 29;148(9):e9-e119. [doi: 10.1161/CIR.0000000000001168] [Medline: 37471501]
2. Bjarnason-Wehrens B, McGee H, Zwisler AD, et al. Cardiac rehabilitation in Europe: results from the European Cardiac Rehabilitation Inventory Survey. Eur J Cardiovasc Prev Rehabil 2010 Aug;17(4):410-418. [doi: 10.1097/HJR.0b013e328334f42d] [Medline: 20300001]
3. Kabboul NN, Tomlinson G, Francis TA, et al. Comparative effectiveness of the core components of cardiac rehabilitation on mortality and morbidity: a systematic review and network meta-analysis. J Clin Med 2018 Dec 4;7(12):514. [doi: 10.3390/jcm7120514] [Medline: 30518047]
4. McCarthy S, O'Raghallaigh P, Woodworth S, Lim YY, Kenny LC, Adam F. Embedding the pillars of quality in health information technology solutions using "Integrated Patient Journey Mapping" (IPJM): case study. JMIR Hum Factors 2020 Sep 17;7(3):e17416. [doi: 10.2196/17416] [Medline: 32940610]
5. Van Schalkwijk DL, Widdershoven J, Elias-Smale S, et al. ShareHeart: a patient journey map of patients with ischemia and non-obstructive coronary artery disease based on qualitative research. J Clin Nurs 2023 Jul;32(13-14):3434-3444. [doi: 10.1111/jocn.16409] [Medline: 35689371]
6. Höppchen I, Kulnik ST, Reich B, et al. Establishing foundations on both sides of the bridge: exploring contextual barriers to cardiac rehabilitation uptake to inform digital health technology design. Presented at: MuC '24 Proceedings of Mensch und Computer 2024; Sep 1-4, 2024; Karlsruhe, Germany. [doi: 10.1145/3670653.3677479]
7. Höppchen I, Wurhofer D, Meschtscherjakov A, Smeddinck JD, Kulnik ST. Targeting behavioral factors with digital health and shared decision-making to promote cardiac rehabilitation-a narrative review. Front Digit Health 2024;6:1324544. [doi: 10.3389/fdgth.2024.1324544] [Medline: 38463944]
8. Höppchen I, Kulnik ST, Meschtscherjakov A. "Be with me and stay with me": insights from co-designing a digital companion to support patients transitioning from hospital to cardiac rehabilitation. Presented at: DIS '24: Proceedings of the 2024 ACM Designing Interactive Systems Conference; Jul 1-5, 2024; Copenhagen, Denmark. [doi: 10.1145/3643834.3661633]
9. Zwack CC, Smith C, Poulsen V, Raffoul N, Redfern J. Information needs and communication strategies for people with coronary heart disease: a scoping review. Int J Environ Res Public Health 2023 Jan 17;20(3):1723. [doi: 10.3390/ijerph20031723] [Medline: 36767091]
10. Weibel L, Massarotto P, Hediger H, Mahrer-Imhof R. Early education and counselling of patients with acute coronary syndrome. a pilot study for a randomized controlled trial. Eur J Cardiovasc Nurs 2016 Jun;15(4):213-222. [doi: 10.1177/1474515114556713] [Medline: 25341680]
11. Greco A, Cappelletti ER, Monzani D, et al. A longitudinal study on the information needs and preferences of patients after an acute coronary syndrome. BMC Fam Pract 2016 Sep 20;17(1):136. [doi: 10.1186/s12875-016-0534-8] [Medline: 27646507]
12. Mitoff PR, Wesolowski M, Abramson BL, Grace SL. Patient-provider communication regarding referral to cardiac rehabilitation. Rehabil Nurs 2005;30(4):140-146. [doi: 10.1097/rnj.0000000000000279]
13. Miller AD, Mishra SR, Kendall L, Haldar S, Pollack AH, Pratt W. Partners in care: design considerations for caregivers and patients during a hospital stay. Presented at: CSCW '16: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; Feb 27 to Mar 2, 2016; San Francisco, CA, USA. [doi: 10.1145/2818048.2819983]
14. Barnason S, Zimmerman L, Nieveen J, Schulz P, Young L. Patient recovery and transitions after hospitalization for acute cardiac events: an integrative review. J Cardiovasc Nurs 2012;27(2):175-191. [doi: 10.1097/JCN.0b013e318239f5f5] [Medline: 22210146]
15. Burgess ER, Reddy MC, Davenport A, Laboi P, Blandford A. Tricky to get your head around": information work of people managing chronic kidney disease in the UK. Presented at: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing System; Glasgow, Scotland, UK. [doi: 10.1145/3290605.3300895]
16. Bickmore TW, Pfeifer LM, Jack BW. Taking the time to care: empowering low health literacy hospital patients with virtual nurse agents. Presented at: CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Boston, MA, USA. [doi: 10.1145/1518701.1518891]
17. Bohanec M, Tartarisco G, Marino F, et al. HeartMan DSS: a decision support system for self-management of congestive heart failure. Expert Syst Appl 2021 Dec;186:115688. [doi: 10.1016/j.eswa.2021.115688]
18. Jacobs ML, Clawson J, Mynatt ED. My journey compass: a preliminary investigation of a mobile tool for cancer patients. Presented at: CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 26 to May 1, 2014; Toronto, Canada. [doi: 10.1145/2556288.2557194]

19. Pfeifer Vardoulakis L, Karlson A, Morris D, Smith G, Gatewood J, Tan D. Using mobile phones to present medical information to hospital patients. Presented at: CHI '12 CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 5-10, 2012; Austin, TX, USA. [doi: 10.1145/2207676.2208601]
20. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, Van Spall HGC. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74. [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]
21. Stamm-Balderjahn S, Falinski R, Rossek S, Spyra K. Development and evaluation of a patient passport to promote self-management in patients with heart diseases. BMC Health Serv Res 2019 Oct 21;19(1):716. [doi: 10.1186/s12913-019-4565-4] [Medline: 31639002]
22. Imran TF, Wang N, Zombeck S, Balady GJ. Mobile technology improves adherence to cardiac rehabilitation: a propensity score–matched study. J Am Heart Assoc 2021 Aug 3;10(15):e020482. [doi: 10.1161/JAHA.120.020482] [Medline: 34278801]
23. Ologeanu-Taddei R. Assessment of mHealth interventions: need for new studies, methods, and guidelines for study designs. JMIR Med Inform 2020 Nov 18;8(11):e21874. [doi: 10.2196/21874] [Medline: 33206060]
24. Wannheden C, Stenfors T, Stenling A, von Thiele Schwarz U. Satisfied or frustrated? A qualitative analysis of need satisfying and need frustrating experiences of engaging with digital health technology in chronic care. Front Public Health 2020;8:623773. [doi: 10.3389/fpubh.2020.623773] [Medline: 33585391]
25. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
26. Hoeppchen I, Kulnik ST, Meschtscherjakov A, et al. Co-designing an information map guiding patients on their journey to cardiac rehabilitation in Austria. Eur J Cardiovasc Nurs 2024 Jul 17;23(Supplement_1):zvae098. [doi: 10.1093/eurjcn/zvae098.073]
27. Sassenrath C, Greving H, Sassenberg K. Are you concerned? Patient testimonials in medical communication affect healthy recipients' emotions and memory. Front Commun 2018 Aug 10;3. [doi: 10.3389/fcomm.2018.00018]
28. Sillence E, Hardy C, Harris PR, Briggs P. Modeling patient engagement in peer-to-peer healthcare. Presented at: WWW '14 Companion: Proceedings of the 23rd International Conference on World Wide Web; Apr 7-11, 2014; Seoul, South Korea. [doi: 10.1145/2567948.2576937]
29. Hassenzahl M, Burmester M, Koller F. AttrakDiff: ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität [Book in German]. In: Szwillus G, Ziegler J, editors. Mensch Comput 2003: Vieweg+Teubner Verlag; 2003:187-196. [doi: 10.1007/978-3-322-80058-9_19]
30. Harborth D, Pape S. German Translation of the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) Questionnaire. SSRN Journal 2018. [doi: 10.2139/ssrn.3147708]
31. Mayring P, Fenzl T. Qualitative inhaltsanalyse [Book in German]. In: Baur N, Blasius J, editors. Handb Methoden Empirischen Sozialforschung: Springer Fachmedien Wiesbaden; 2019:633-648. [doi: 10.1007/978-3-658-21308-4_42]
32. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006 Mar;25(2):91-97. [doi: 10.1080/01449290500330331]
33. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367. [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
34. Liu H, Wilton SB, Southern DA, et al. Automated referral to cardiac rehabilitation after coronary artery bypass grafting is associated with modest improvement in program completion. Can J Cardiol 2019 Nov;35(11):1491-1498. [doi: 10.1016/j.cjca.2019.08.008] [Medline: 31604671]
35. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015 Feb 12;10(1):21. [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
36. Pratt W, Unruh K, Civan A, Skeels MM. Personal health information management. Commun ACM 2006 Jan;49(1):51-55. [doi: 10.1145/1107458.1107490]
37. Philip B, Abdelrazek M, Barnett S, Bonti A, Grundy J. Towards better mHealth apps: understanding current challenges and user expectations. Presented at: MOBILESoft '22: IEEE/ACM 9th International Conference on Mobile Software Engineering and Systems; May 17-24, 2022; Pittsburgh, PA, USA. [doi: 10.1145/3524613.3527804]
38. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
39. Schneider-Matyka D, Wojcik J, Rachubinska K, Grochans E, Cybulska AM. Perceived social support and stress-coping strategies among cardiac patients. Eur J Cardiovasc Nurs 2024 Jul 17;23(Supplement_1):104. [doi: 10.1093/eurjcn/zvae098.104]
40. Fortuna KL, Brooks JM, Umucu E, Walker R, Chow PI. Peer support: a human factor to enhance engagement in digital health behavior change interventions. J Technol Behav Sci 2019 Jun 15;4(2):152-161. [doi: 10.1007/s41347-019-00105-x] [Medline: 34337145]

41. Parry M, Watt-Watson J. Peer support intervention trials for individuals with heart disease: a systematic review. Eur J Cardiovasc Nurs 2010 Mar;9(1):57-67. [doi: 10.1016/j.ejcnurse.2009.10.002] [Medline: 19926339]
42. Flobak E, Nordberg OE, Guribye F, Nordgreen T, Sekse RJT. "This is the story of me": designing audiovisual narratives to support reflection on cancer journeys. Presented at: DIS '21: Proceedings of the 2021 ACM Designing Interactive Systems Conference; Jun 28 to Jul 2, 2021. [doi: 10.1145/3461778.3462005]
43. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. Field methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
44. Waddell A, Seguin JP, Wu L, et al. Leveraging implementation science in human-centred design for digital health. Presented at: CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems; May 11-16, 2024; Honolulu, HI, USA. [doi: 10.1145/3613904.3642161]
45. Bonawitz K, Wetmore M, Heisler M, et al. Champions in context: which attributes matter for change efforts in healthcare? Implement Sci 2020 Aug 6;15(1):62. [doi: 10.1186/s13012-020-01024-9] [Medline: 32762726]
46. Scott S, Brameier DT, Tryggedsson I, et al. Prevalence, resources, provider insights, and outcomes: a review of patient mental health in orthopaedic trauma. J Orthop Surg Res 2024 Sep 3;19(1):538. [doi: 10.1186/s13018-024-04932-4] [Medline: 39223649]

## Abbreviations

**CR:** cardiac rehabilitation
**mHealth:** mobile health
**NASSS :** Nonadoption, Abandonment, Scale-Up, Spread and Sustainability
**UTAUT2:** Unified Theory of Acceptance and Use of Technology 2

*Edited by G Krstačić; submitted 16.04.25; peer-reviewed by A Müller, D Azar, V Janssen; revised version received 17.07.25; accepted 21.07.25; published 27.08.25.*

*Please cite as:*
*Höppchen I, Kulnik ST, Meschtscherjakov A, Niebauer J, Reich B, Smeddinck JD, Wurhofer D*
*mHealth Support in Cardiac Care Pathways for Patient Self-Management During Transitions From Hospital to Rehabilitation: Exploratory Field Study*
*JMIR Cardio 2025;9:e76089*
*URL: https://cardio.jmir.org/2025/1/e76089*
*doi:10.2196/76089*

# Health Care Professionals' Use of Digital Technology in the Secondary Prevention of Cardiovascular Disease in Austria: Online Survey Study

Luisa Lunz[1,2,3], MSc; Sabine Würth[3], PhD; Stefan Tino Kulnik[1], MRes, PhD

[1]Ludwig Boltzmann Institute for Digital Health and Prevention, Lindhofstrasse 22, Salzburg, Austria

[2]Department of Health Sciences, Salzburg University of Applied Sciences, Salzburg, Austria

[3]Department of Sport and Exercise Science, University of Salzburg, Salzburg, Austria

**Corresponding Author:**
Stefan Tino Kulnik, MRes, PhD
Ludwig Boltzmann Institute for Digital Health and Prevention, Lindhofstrasse 22, Salzburg, Austria

## Abstract

**Background:** Advances in digital technology, such as health apps and telerehabilitation systems, offer promising treatment modalities in the secondary prevention of cardiovascular disease. However, the successful adoption of digital technology in clinical practice depends on a variety of factors. A comprehensive understanding of the influencing factors on digital technology usage in health care can support the complex implementation process of digital technology in clinical practice.

**Objective:** The aim of this study was to identify barriers and facilitators of digital technology usage in cardiovascular disease secondary prevention from the perspective of health care professionals, and to explore whether certain characteristics of health care professionals are related to the current usage of digital technology in clinical practice.

**Methods:** We conducted an exploratory online survey, inquiring about the perspectives and uses of digital technologies in cardiovascular disease secondary prevention. We developed an original questionnaire to address the study aim. The survey invitation was distributed among health care professionals from November 2021 to February 2022, via all cardiac rehabilitation centers, all community-based disease management services for patients with chronic heart failure, and all relevant national health care professional associations in Austria. Qualitative survey data were analyzed using thematic content analysis. Quantitative survey data were analyzed using descriptive statistics, group comparison tests, and association statistics.

**Results:** Overall, 125 health care professionals (mean age 41, SD 11 y; n=80, 64% females) across different professions and settings, including cardiac rehabilitation phases I through IV, were recruited. General readiness for using digital technologies in the care of cardiac patients was high, but only 65 (52%) respondents reported doing so. The top 3 rated barriers to digital technology use were poor user-experience of devices and apps, lack of cost coverage, and low digital competence of patients. The top 3 rated potential application areas for digital technology were organization and appointment planning, documenting treatments, and creating personalized treatment plans. The top 3 rated facilitators for digital technology use were assurance of patient safety, assurance of patients' privacy, and availability of technical support. Greater personal use of digital technology, younger age, and higher technology affinity of health care professionals was associated with higher readiness to use digital technology with cardiac patients.

**Conclusions:** While there is interest in digital technology for the secondary prevention of cardiovascular disease in Austria, barriers to uptake need to be addressed. Our findings may inform the design and implementation of future digitalization projects.

***(JMIR Cardio 2025;9:e71366)*** doi:10.2196/71366

**KEYWORDS**

barriers; cardiac rehabilitation; cross-sectional; digital health; electronic health; facilitators; mobile health; questionnaire; telerehabilitation

## Introduction

Cardiovascular disease (CVD) remains the leading cause of death and a large contributor to loss of healthy life expectancy worldwide [1,2]. The modification of cardiovascular risk factors can have a positive influence on reducing this burden and has been a main focus of secondary prevention, for example, through exercise-based cardiac rehabilitation (CR) programs [3,4]. However, this assumes that patients can consistently adopt heart-healthy behavior changes into their daily lives, which often poses a major challenge [5].

Advances in digital technology (DT) are opening up promising ways to help patients change and self-manage their lifestyle [6]. For example, the recent European Society of Cardiology guidelines for the management of chronic coronary syndromes now include a class 1A recommendation for mobile health interventions to improve patient adherence to healthy lifestyles and medical therapy [7]. Such interventions, incorporating text messaging, smartphone apps, web-based content, and wearable devices, have been shown to support patients' healthy behaviors including medication adherence [8], exercise habits [8-11], and diet [9]. Demonstrated effects on clinical outcomes are improved blood pressure control [8,10], increased exercise capacity [12], reduced waist circumference [10], reduced low-density lipoprotein levels [9,10], decreased incidence of major adverse cardiovascular events [12], and improved quality of life [12].

Furthermore, DT has facilitated the provision of telerehabilitation, that is, home-based CR programs delivered remotely by CR professionals, which could increase access to a structured and supervised exercise-based CR program for patients who are unable or unwilling to attend a center-based CR program [13]. High-level evidence shows that telerehabilitation compared to center-based CR offers equivalent effects on patient outcomes in terms of medication adherence, smoking behavior, physiological risk factors, depression, functional capacity, exercise behavior, cardiac-related hospitalization, and quality of life [14,15].

While the scientific evidence for DT in the secondary prevention of CVD is strong, its implementation in real-life practice often lags behind [16]. The successful adoption of DT in clinical practice depends on a variety of factors, for instance, on the technology itself, its promised benefits for patients, organizational and systematic factors, as well as the characteristics, attitudes and experiences of the various user-groups (eg, patients, caregivers, and health care professionals [HCPs]) [16,17]. The scoping review by Whitelaw at al [18], for example, lists the following commonly reported clinician-level barriers to uptake of DT in cardiovascular care: increased work and responsibilities, unreliable technologies, lack of evidence supporting the use of technology, and lack of integration with medical records. The most commonly described clinician-level facilitators were approval and organizational support from senior management and improved efficiency through DT [18]. Because the organization, structure, and funding of health care systems can differ considerably from country to country, a comprehensive understanding of the influencing factors on DT usage in a national health care context can support the complex implementation process of DT in clinical practice [17].

The aim of this study was to identify barriers and facilitators of DT usage in CVD secondary prevention from the perspectives of HCPs in Austria. Specifically, we sought to identify HCPs' attitudes toward DT usage, and to explore whether certain HCP characteristics (affinity for DT, personal use of DT, age, physical activity [PA] behavior, and professional background) are related to the current usage of DT in clinical practice.

## Methods

### Overview

We conducted a cross-sectional online survey among HCPs working in the secondary prevention of CVD in Austria. In the reporting of this study, we adhere to the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [19].

### Setting and Participants

Our survey addressed settings for the secondary prevention of CVD in Austria, including general practitioner and cardiologist practices, outpatient clinics, community-based disease management programs for patients with chronic heart failure, and the CR pathway. In Austria, the latter comprises 4 phases: the acute hospital stay (phase I), medically supervised in- or outpatient rehabilitation programs of up to 6 weeks duration (phase II), medically supervised outpatient rehabilitation programs of 6 - 12 months duration with weekly or less frequent sessions (phase III), and patients' life-long independent secondary prevention behavior and self-management (phase IV) [20]. We invited qualified HCPs from any relevant professional background (including nurses, physicians, sport scientists, physiotherapists, psychologists, and dietitians) who were working in any of these settings. Unemployed HCPs and retirees were excluded from the survey.

### Recruitment

Recruitment took place between November 1, 2021, and February 20, 2022. Email invitations with an open link to the online questionnaire were sent to the medical and nursing directors of all CR centers (at the time 13 inpatient and 21 outpatient centers); to all 3 community-based disease management services for patients with chronic heart failure; and to the boards of all relevant HCP associations (cardiology, dietetics, nursing, nutrition science, occupational therapy, physiotherapy, psychology, social work, and sports science) in Austria. The addressees were asked to forward the survey invitation to all employees or members of their organizations.

### Sample Size

This exploratory survey recruited a convenience sample, and no prospective sample size calculation was conducted. Based on response rates from previous online surveys among HCPs in Austria that used similar recruitment strategies, we expected to achieve a sample size of 100 to 200 respondents.

### Survey Instrument Development

Because no valid survey instrument existed that aligned with the study aims, an original questionnaire was designed, implemented in LimeSurvey (version 3.25.21+210407) and piloted. The questionnaire's content was developed based on qualitative data (interviews and focus group) from 7 CR professionals and literature on obstacles and potential application areas for DT in health care [21-23]. Then, the questionnaire was piloted using cognitive interviewing with 8 HCPs from different professional backgrounds who were representative of the target sample. The questionnaire was iterated and revised twice to optimize comprehensibility, usability, completion rate, and completion time. The development process supports content

and construct validity of the survey instrument, but we were unable to perform psychometric assessments of construct validity (eg, convergent validity) due to the lack of suitable validated alternative measures.

### Questionnaire

The questionnaire consisted of 42 items divided into 10 sections. Items were formulated as multiple-choice questions, Likert scale items [24] and open questions with free-text answers. The estimated completion time was 20 minutes. The full questionnaire in its original German version is available at the Open Science Framework platform [25], and an English translation is given in Multimedia Appendix 1. In summary, the questionnaire covered the following content:

- Professional profile (5 items).
- PA behavior (meeting the World Health Organization [WHO] recommendations [26]; 4 items).
- Affinity for DT (2 polarizing questions selected from the TA-EG questionnaire [27]; 1 item).
- Personal use of DT (types of digital devices used—in particular devices relating to PA, 3 items).
- Use of DT at work in cardiovascular care (recommending the use of DT to patients, reasons for recommending or not recommending DT to patients, types of DT used with patients or for patient care, reasons for non-use of DT, DT used for certain patient groups only, past use of DT and reasons for discontinuing, knowledge of DT used in cardiovascular care by other HCPs or in another setting; 12 items).
- Readiness for using DT in their work (1 item).
- Perceived barriers to using DT in cardiovascular care (rating of 20 potential barriers, 1 open question; 5 items).
- Potential application areas for DT in cardiovascular care (rating of 17 potential application areas, 1 open question; 3 items).
- Factors influencing the decision to use or not use DT (rating of 22 potential influencing factors, 1 open question; 4 items).
- Demographic information (gender, age, highest education level, professional qualification; 4 items).

### Ethical Considerations

The study was reviewed by the research ethics committee of the University of Salzburg and received favorable ethical opinion (reference GZ 21/2021). Survey respondents were presented with information about the study and contact details of the study team on the first page of the online questionnaire. Respondents had to first confirm their informed consent in the online questionnaire, before anonymously completing the survey questions. A voluntary prize draw of three smart watches and fitness trackers, worth US $200 each, served as incentive for participation. To maintain anonymity of survey responses, email addresses required for prize notification were collected separate from the survey responses.

### Data Cleaning

Verification of data completeness was not necessary, as only fully completed surveys were saved to the platform. Individual respondents' completion time was reviewed to reduce the likelihood of dishonest answers (eg, overly fast completion time).

### Data Analysis

Qualitative data from free-text answers were analyzed using thematic content analysis [28]. Quantitative data were analyzed descriptively. Group comparisons were conducted using *t* test, Man-Whitney *U* test and Kruskal-Wallis test (2-tailed, alpha=.05). We calculated associations between HCP's PA behavior and their personal use of DT, and HCP's age, sex, and professional background and their affinity for DT. To explore whether certain HCP characteristics were related to the current use of DT in clinical practice, we calculated bivariate association statistics between the predictor variables affinity for DT, personal use of DT, age, sex, PA behavior, and professional background, and the outcome variables DT recommendation behavior, DT implementation behavior and readiness to use DT in practice, applying the appropriate statistical tests ($\chi^2$ test, binary logistic regression, and Spearman correlation coefficient). All statistical analyses were performed in SPSS software (version 22.0, IBM) and without correction for multiple testing due to their purely exploratory nature.

## Results

The survey recruited 125 participants. Respondent characteristics are given in Table 1.

**Table .** Respondent characteristics.

| Characteristic | Sample (N=125), n (%) |
|---|---|
| Age (years), mean (SD) | 41 (11) |
| Sex | |
| Female | 80 (64) |
| Male | 38 (30) |
| Nonbinary | 1 (1) |
| Not disclosed | 6 (5) |
| Education | |
| Compulsory schooling, apprenticeship | 8 (7) |
| A-levels or equivalent professional education | 24 (19) |
| University | 93 (74) |
| Professional qualification[a] | |
| Nursing | 40 (32) |
| Medicine | 25 (20) |
| Sports science | 21 (17) |
| Other | 17 (14) |
| Physiotherapy | 13 (10) |
| Psychology | 13 (10) |
| Dietetics | 10 (8) |
| Medical assistant | 3 (2) |
| Administration | 1 (1) |
| Social work | 1 (1) |
| Clinical remit[a] | |
| Nursing care | 40 (32) |
| Medical care | 25 (20) |
| Medical training therapy | 23 (18) |
| Administration | 19 (15) |
| Social work | 19 (15) |
| Physiotherapy | 12 (10) |
| Nutrition advice | 11 (9) |
| Psychological care | 9 (7) |
| Smoking cessation | 7 (6) |
| Other | 5 (4) |
| Sports science | 2 (2) |
| Setting[a] | |
| Outpatient rehabilitation center | 43 (34) |
| Inpatient rehabilitation center | 41 (33) |
| Acute hospital – inpatients | 29 (23) |
| Private practice | 12 (10) |
| Acute hospital – outpatients | 8 (6) |
| Patient home visits | 5 (4) |
| Other | 4 (3) |

| Characteristic | Sample (N=125), n (%) |
|---|---|
| Non–health care setting | 1 (1) |
| Cardiac rehabilitation phase[a] | |
| Phase I | 27 (22) |
| Phase II | 79 (63) |
| Phase III | 44 (35) |
| Phase IV | 33 (26) |
| Community-based disease management program for patients with chronic heart failure | 8 (6) |
| Other | 5 (4) |

[a]Multiple answers possible.

## HCPs' Affinity for DT, Personal Use of DT and PA behavior

Most HCPs rated themselves tech-savvy (median 2, IQR 2-3; on a 5-point Likert scale from 1 ["very tech-savvy"] to 5 ["not at all tech-savvy"]). Only 5 (4%) respondents reported no personal use of DT, with others using smartphones (n=114, 91%), wrist-worn heart rate sensors (n=51, 41%), smartwatches (n=35, 28%), step counters (n=33, 26%), watches with chest strap for heart rate measurement (n=27, 22%), and digital devices for measuring physical performance (n=12, 10%). Older HCPs had lower affinity for use of DT (rho=0.24, 95% CI 0.06 - 0.41; *P*=.006). There were no significant differences in affinity according to sex or professional group. A total of 54 (43%) respondents reported meeting the WHO PA recommendations for adults (≥150 minutes per week of moderate or ≥75 minutes per week of vigorous intensity endurance-type PA; and ≥2 times per week muscle strengthening activities) [26], with 56 (45%) reporting recording, planning or sharing their PA using DT. Binary logistic regression revealed a higher likelihood of personal use of DT (in particular devices with PA-related functionalities) for those who met the PA recommendations, as compared to those who did not (OR 2.8, 95% CI 1.4 - 5.9; *P*=.005).

## Recommendation and Usage of DT in Practice

Respondents' subjective readiness to use DT in clinical practice was high (median 2, IQR 1-2; on a 5-point Likert scale from 1 ["very inclined"] to 5 ["very opposed"]). Overall, 88 (70%) respondents reported that they currently recommended the use of DT to their CVD patients. A total of 80 respondents listed their most common recommendations in free text answers. These were for smartwatches and heart rate monitors (n=47, 59%), apps (n=36, 45%), and step counters (n=16, 20%), primarily for aspects of training control, heart rate monitoring, and recording or visualizing of vital signs, training and PA behavior. A total of 65 (52%) HCPs reported currently using DT as part of their clinical practice with CVD patients, including chest straps (n=32, 49%) and wrist watches (n=17, 26%) for heart rate measurement, apps (n=21, 32%), online information (n=12, 18%), step counters (n=12, 18%), smartwatches (n=11, 17%), and activity trackers (n=10, 15%). The most used apps were HerzMobil, heartfish, and RehaApp. HerzMobil (Landesinstitut für Integrierte Versorgung Tirol, Innsbruck, Austria) is part of a telemonitoring system in conjunction with Bluetooth-enabled blood pressure devices and scales [29]. The system was offered by one regional heart failure disease management service in Austria. The cost of HerzMobil was covered by a regional public healthcare fund. heartfish (heartfish GmbH, Vienna, Austria) is an app that aims to support motivation and adherence with exercise therapy in patients with CVD, cancer, and other conditions [30]. heartfish was in use at several outpatient CR centers in Austria. The basic version of the app was made available to patients for free, with the option of a paid subscription for extended functionalities. RehaApp (Pensionsversicherungsanstalt, Vienna, Austria) is an app to support self-monitoring of blood pressure, body weight, medication, and PA adherence following an inpatient rehabilitation stay. The app was in use as part of a clinical trial at inpatient CR centers in Austria.

## Reasons for Non-Recommendation and Non-Usage of DT in Practice

Reasons given in free text for non-recommendation and non-usage of DT in practice are listed in Table 2. The most common reasons for not recommending DT to CVD patients were the feeling of it not being within one's area of responsibilities or allocated tasks, lacking technical skills, as well as concerns over the patient becoming too dependent on DT or reducing their sense of body awareness. The most frequently given reason for not using DT in practice was lack of opportunity or possibility to do so, followed by the patient's age, not feeling responsible for it, lack of familiarity with suitable options, and not having enough time.

**Table .** Reasons for non-recommendation and non-usage of digital technologies in practice.

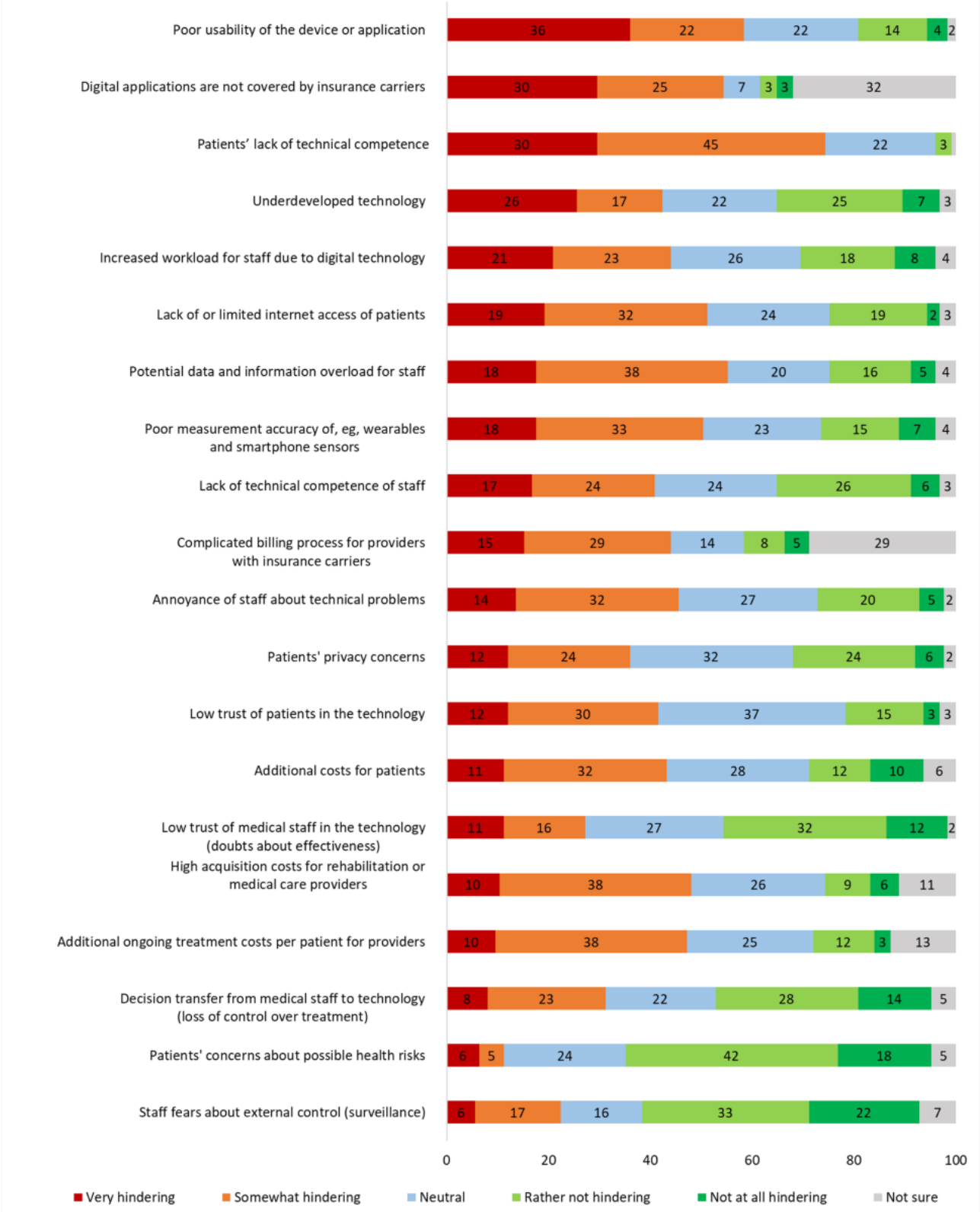
| Question | Responses | | |
|---|---|---|---|
| | Relating to the patient | Relating to the health care professional | Relating to the physical and social environment |
| If you can think of any specific reasons why you do not recommend digital technologies to your patients, please describe them here (n=21) | • Concerns regarding loss of body awareness and risk of dependence on digital technologies (n=4)<br>• Too overwhelming (n=3)<br>• Age (n=2)<br>• Pressure to perform (n=2)<br>• Compliance (n=1) | • Not within one's own area of responsibility or tasks (n=5)<br>• Lack of own technical competence (n=4)<br>• Lack of exposure to possible digital technology (n=3)<br>• Lack of time (n=2)<br>• Not interested in advertising products (n=1) | —[a] |
| If you can think of any specific reasons why you do not implement digital technologies into your clinical practice, please describe them here (n=30) | • Age (n=5)<br>• Patients already use digital tools independently (n=1) | • Perceived as outside one's responsibility (n=3)<br>• Lack of familiarity with practical, appropriate, ad-free options (n=3)<br>• Lack of time (n=3)<br>• Focus on personal coaching (n=1)<br>• Lack of communication skills (n=1) | • Lack of opportunity or possibility (n=11)<br>• Poor internet connection (n=1)<br>• Lack of implementation in the work process (n=1) |

[a]Not applicable.

## Barriers

The top 5 rated barriers (answer "very hindering") of using DT in practice were poor usability, lack of reimbursement from insurance carriers, patients' lack of technical competence, underdeveloped technology, and fear of increased workload for staff (see Figure 1). The latter point was reiterated 8 times in the free-text answers.

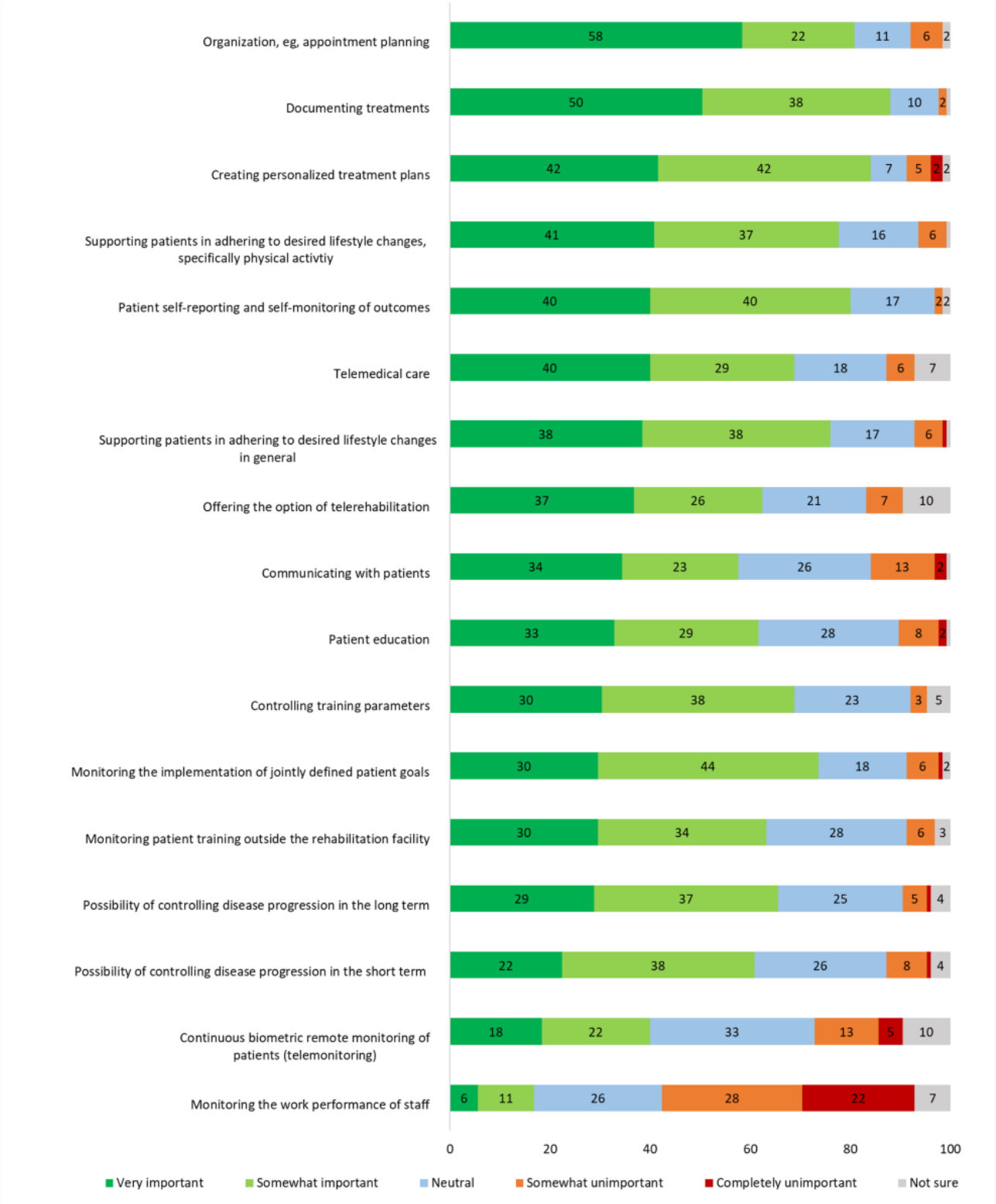
**Figure 1.** Barriers to the use of digital technologies in the secondary prevention of cardiovascular disease. Respondents (N=125) rated each potential barrier on a 5-point Likert scale. The percentages for each response category are shown.




## Potential Areas of Application

The application areas for DT that were perceived as most relevant (marked "very important") were for organization, documentation of measurements, creating personalized treatment plans, supporting patients in their adherence to PA lifestyle change, and patient self-reporting of outcomes (see Figure 2).

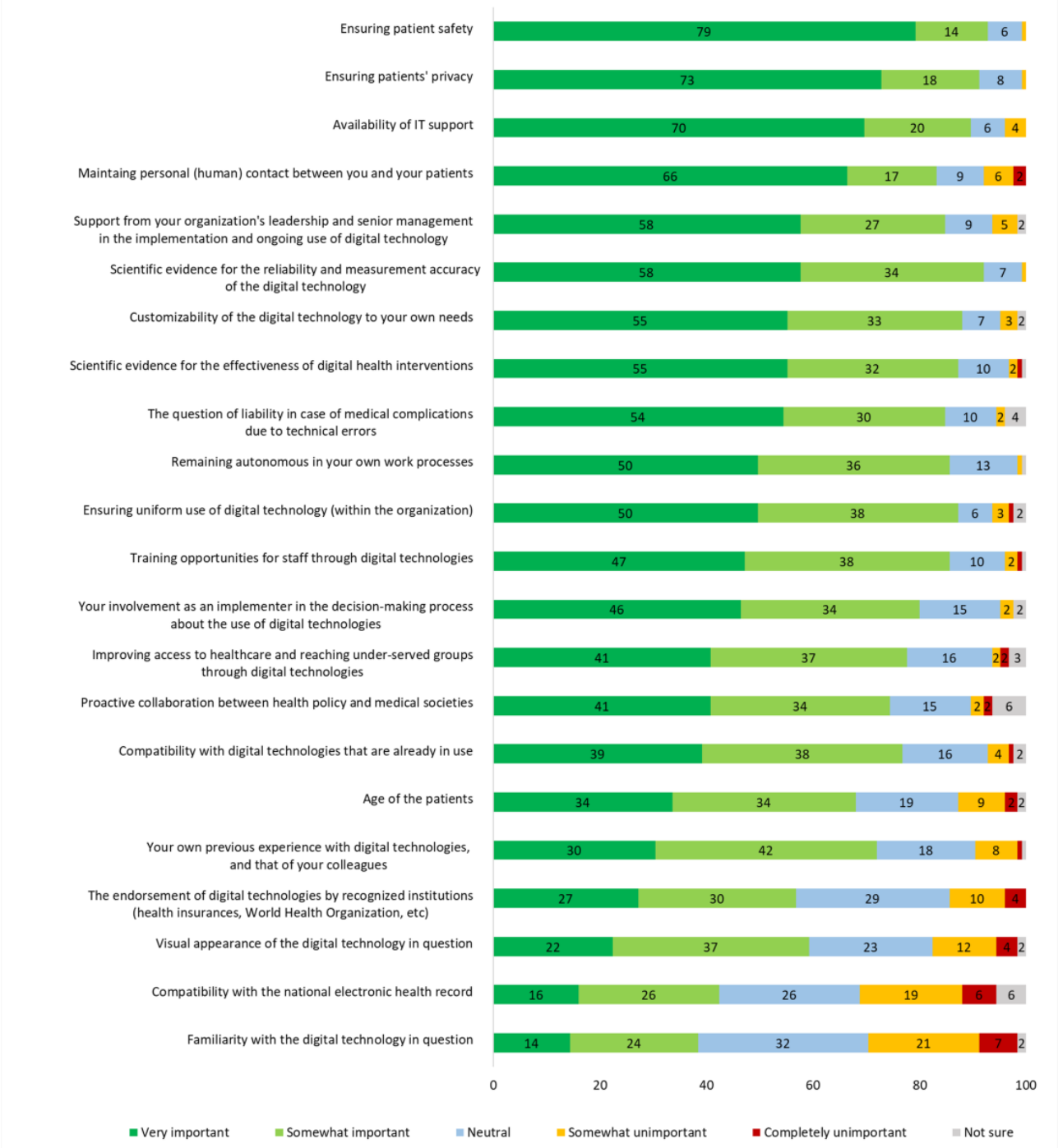
**Figure 2.** Potential important application areas of digital technologies in the secondary prevention of cardiovascular disease. Respondents (N=125) rated each potential application area on a 5-point Likert scale. The percentages for each response category are shown.




## Factors Influencing the Decision to Use DT

The highest rated influencing factors (marked “very important”) for using DT in practice were assurance of patient safety and privacy, availability of technical support, and the maintenance of personal contact between HCPs and their patients (see Figure 3).

**Figure 3.** Factors that influence the decision to use or not use digital technologies in the secondary prevention of cardiovascular disease. Respondents (N=125) rated each potential factor on a 5-point Likert scale. The percentages for each response category are shown.




## Recommending DT to Patients

HCPs' inclination to recommend DTs to their CVD patients was not related to HCPs' own PA behavior or their personal use of DT, nor did the mean age of those who recommended DT (41, SD 12 y) versus those who did not (40, SD 10 y) differ significantly. In addition, affinity for DT did not significantly differ for recommenders and non-recommenders, with a median of 2 (IQR 1.5-3; "rather tech-savvy") in both groups. However, the likelihood of recommending DTs was significantly higher among medical doctors compared to other professions (OR 7.3, 95% CI 1.4 - 38.3; *P*=.02).

## Implementing DT in Clinical Practice With Patients

The use of DT in clinical practice was not statistically related to HCP's own PA behavior, their personal use of DT, their sex or professional and academic background, nor did the mean age of users (42, SD 12 y) and non-users (39, SD 10 y) or affinity for DT across users and non-users differ significantly.

## Readiness to Use DT in Practice

HCP's subjective readiness to use DT in practice was not related to their own PA behavior or professional and academic background. However, descriptively, sport scientists reported the highest readiness to use DT with a median of 1.5 (IQR 1-2),

and psychologists the lowest with a median of 2.5 (IQR 2-3). Respondents who personally used DT demonstrated a significantly higher readiness to do so in clinical practice, as compared to those who did not (mean 1.7 SD 0.8 vs 2.2 SD 1.0, respectively; *P*<.001). Furthermore, older HCPs felt less ready to use DTs in practice ($r_s$=0.22, 95% CI 0.04 - 0.39; *P*=.01), and those with higher affinity for DT felt more ready to use DTs in practice ($r_s$=0.47, 95% CI 0.31 - 0.60; *P*<.001).

## Discussion

### Principal Findings

We found that respondents' readiness and attitudes toward using DT in the secondary prevention of CVD were generally positive. However, in comparison, their current usage of DT in practice was relatively low at just over 50% across all professions, and particularly low among dieticians, nurses, physiotherapists, and psychologists, of whom less than half reported implementing DT.

HCPs' age was not significantly related to the usage or recommendation of DT in clinical practice, but older age was associated with lower readiness for DT implementation and lower affinity for DT. The threat of ageism to successful digital engagement is increasingly being highlighted, and authors call for awareness-raising and training to achieve a positive framing of older age in the digital world [31]. As such, HCPs' age may not constitute a major obstacle in the usage of DT but should be considered in the training and integration phases of DT in clinical practice. The successful implementation of new DT requires organizational and collegial support [32]. For instance, specialized training options offered to older HCPs might increase readiness to use DT, and thus, contribute to a successful implementation of DT in clinical practice. Furthermore, peers who act as implementation "champions" can assist in building positive experiences of digitalization for their colleagues. The concept of implementation champions stems from implementation science and describes a role occupied by people who are internal to an organization, have an intrinsic interest to implementing a change, and are committed to drive implementation forward [33]. Our data describe a profile of younger, more physically active HCPs with greater affinity and personal use of DT and higher readiness to use DT with patients. Such individuals, among others, could be enlisted to act as implementation champions and peer supporters for colleagues.

In terms of PA, HCPs who met the WHO PA recommendations were nearly 3 times as likely to personally use DT than those who did not. Thus, it is plausible that PA increases with DT usage, as studies report increased daily active minutes and steps through smartphone app or wearable usage [34]. On the other hand, physically active respondents may simply be more inclined to use DT to manage or track their PA, which would be reflective of the types of fitness apps that respondents in our sample most frequently used in their private lives (ie, Strava, Garmin Connect, and Polar Flow). Interestingly, the personal usage of smartwatches for heart rate measurement was less frequently reported by respondents, with just slightly over half reporting so.

In HCPs' work-related use of DT, heart rate monitors, smartwatches, and apps were the most frequently used and recommended devices. Regarding the named apps, it is noticeable that these target multiple cardiovascular risk factors and clinical parameters. Apps that focus on single health behaviors (eg, smoking cessation) or more specific clinical issues (eg, mental health) were not listed. There is some evidence to suggest that digital health interventions that target multiple health behaviors or CVD risk factors could be more potent, for example, in the systematic review by Akinosun et al [9]. But apps that focus on single health behaviors or cross-cutting topics such as mental health could be equally relevant and appropriate in CVD secondary prevention [35], and such apps are currently more widely available than CVD-specific mobile health solutions, for example, in the German directory of approved and reimbursed digital health applications [36]. The prevalence of chest straps for heart rate measurement was higher than wrist-worn sensors, possibly due to chest straps having been established for longer in CR. But it may also reflect that many wrist-worn sensors are still less accurate than chest straps for measuring heart rate, which would correspond with the eighth-rated barrier in our survey [37].

With regard to potential application areas for DT in the secondary prevention of CVD, HCPs perceived organization and appointment scheduling as the most relevant, especially in the early phases of CR when regular contact and scheduling is required, followed by documentation of treatments. For instance, a uniform, digital system could be helpful in seamlessly tracking measurements. At home, the use of an app could allow patients to visualize results, better inform themselves, and monitor their own parameters. Creating personalized treatment plans and supporting patients with behavioral changes (specifically PA behavior but also desired lifestyle changes in general) were other highly ranked potential application areas, which mirrors other studies of HCP's perceptions of digital health in cardiac care [38]. A further highly ranked potential application area concerns the provision of remote care, including telemedical care in the sense of remote individual consultations via video or telephone calls as well as offering structured and supervised CR programs via telerehabilitation formats in addition to center-based in- or outpatient CR. While the COVID-19 pandemic has to some extent forced HCPs to establish remote formats for individual consultations, telerehabilitation options for phase II or III CR programs are still lacking in Austria to date, despite their potential to increase the reach and uptake of CR among patients who do not engage with center-based rehabilitation [13].

The highest-rated barriers to DT usage in our survey included poor usability, increased workload for staff, patient age, and lack of cost coverage, which corresponds with commonly reported barriers in the literature, for example, in the scoping review of 29 primary studies by Whitelaw et al [18]. In our qualitative survey responses, concerns over patients' dependence on DT was the most frequently listed patient-related barrier, corroborating some smaller qualitative studies, which have also raised this point. For instance, Attig and Franke [39] reported decreased PA motivation when commonly worn fitness trackers were not available for users, for example, when the device had

been forgotten or its battery was empty; and other qualitative studies of CR patients have observed patients' own concerns about dependence on DT [40]. However, the number of studies reporting positive effects of fitness tracking on users' motivation to be physically active [10,12] suggests that, while a risk of dependence should be taken into account, the increased motivation elicited by DT may outweigh the potential consequences of dependence.

Poor usability and increased workload were also reported barriers in a recent qualitative study that evaluated the implementation of a digital CR intervention [41]. Poor usability and increased workload go hand-in-hand, as poor usability increases workload demands. As such, well-designed and optimized DT can aide in overcoming these barriers. User-centered co-design constitutes a methodological cornerstone to achieve this and is gradually finding increasing application in the development of interventions for the secondary prevention of CVD [42].

Old age or perception of age-related barriers, such as DT not being suitable for older patients, were reported hindrances of DT usage in clinical practice. As this can lead to perpetuation of negative ageist stereotypes and exclusion of older patients from digital health interventions [31], consideration of ways to facilitate older patients' participation in DT usage is needed. In addition to individual-level strategies such as communicating personal benefits of DT for older people and offering age-tailored instructional materials and training in DT use to patients [43], meso-level strategies are required, including changing the negative discourse on aging, and inclusion and partnership with older people in the design of DT and digital health care services [31,44]. Rather than gatekeeping the provision of DT according to the perceived digital competency of patients, HCPs may find that many individuals who are less familiar with DT are able to engage with digital health interventions with minimal assistance [45].

Finally, the lack of cost coverage by insurance providers hindered HCPs from using DT. Although, there is good scientific evidence of the health-promoting effects of DT in the secondary prevention of CVD [46,47], there is currently still no established reimbursement system for digital health interventions in Austria and many other European countries. Austria's journey towards embracing digital health started 2 decades ago, with the decision to introduce a national electronic health record system [48]. But concrete efforts towards a reimbursement system for digital health interventions have only started in 2023, concurrently with the development of the first national eHealth strategy for Austria [49]. While other European countries, notably Belgium, Germany, and France, have been more proactive in setting up transparent reimbursement systems for digital health interventions [50], Austria plans to create a process by 2026, which is expected to act as a catalyst for the implementation of DT in clinical practice. In this, it will be important to guard against inherent inequity and widening of the digital divide, which is driven not only by the direct costs of DT to HCPs and patients (eg, licenses and subscriptions), but also by structural and socioeconomic disadvantage among the population, including the lack of network infrastructure (internet broadband access, data allowance), the affordability of smartphones and computers, and limited digital literacy [51,52]. In Austria's publicly funded health care system with near-universal coverage [53], direct costs of DT can be expected to have lesser impact on inequity, but structural and socioeconomic disadvantage alongside collateral and hidden costs for enabling inclusive digital health, such as the provision of digital skills training for patients, need to be taken into account.

## Limitations

Our survey was limited by the self-selected nature of the sample, leading to possible selection bias towards individuals with interest in the topic, for example, those with greater affinity and more positive attitudes towards DT. This likely accounts for the high levels of affinity for DT and subjective readiness to use DT in clinical practice among the sample. We acknowledge that the questionnaire did not capture respondents' responsibility or role with regard to DT in clinical practice, that is, whether they were a prescriber or they executed a prescription. Although we were able to recruit respondents across the different professions involved in CVD secondary prevention in Austria, our findings are to be interpreted as exploratory rather than representative. The lack of a prospective sample size calculation is acknowledged.

## Conclusion

We conducted the first nationwide Austrian survey to capture HCPs' perspectives and use of DT in CVD secondary prevention. We describe the currently prevalent types of digital health interventions and digital devices and give insight into HCPs' perspectives on relevant application areas, barriers, and facilitators for DT in CVD secondary prevention. These findings can sensitize digital intervention developers, researchers, and implementers to HCPs' needs and wants with regard to DT, thereby contributing to the successful design and implementation of digitalization projects in CVD secondary prevention.

## Acknowledgments

We thank all cardiac rehabilitation centers and health care professional associations for disseminating the survey link among their staff and membership. We extend our gratitude to all health care professionals who took part in this survey. We thank Hannah McGowan for the language editing of this manuscript.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
English translation of the online questionnaire.
[DOCX File, 43 KB - cardio_v9i1e71366_app1.docx ]

## References


1. Roth GA, Mensah GA, Fuster V. The global burden of cardiovascular diseases and risks: a compass for global action. J Am Coll Cardiol 2020 Dec 22;76(25):2980-2981. [doi: 10.1016/j.jacc.2020.11.021] [Medline: 33309174]
2. Vos T. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
3. Visseren FLJ, Mach F, Smulders YM, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur Heart J 2021 Sep 7;42(34):3227-3337. [doi: 10.1093/eurheartj/ehab484] [Medline: 34458905]
4. Laranjo L, Lanas F, Sun MC, et al. World Heart Federation roadmap for secondary prevention of cardiovascular disease: 2023 update. Glob Heart 2024;19(1):8. [doi: 10.5334/gh.1278] [Medline: 38273995]
5. Kotseva K, De Backer G, De Bacquer D, et al. Lifestyle and impact on cardiovascular risk factor control in coronary patients across 27 countries: results from the European Society of Cardiology ESC-EORP EUROASPIRE V registry. Eur J Prev Cardiol 2019 May;26(8):824-835. [doi: 10.1177/2047487318825350] [Medline: 30739508]
6. Frederix I, Caiani EG, Dendale P, et al. ESC e-cardiology working group position paper: overcoming challenges in digital health implementation in cardiovascular medicine. Eur J Prev Cardiol 2019 Jul;26(11):1166-1177. [doi: 10.1177/2047487319832394] [Medline: 30917695]
7. Vrints C, Andreotti F, Koskinas KC, et al. 2024 ESC Guidelines for the management of chronic coronary syndromes. Eur Heart J 2024 Sep 29;45(36):3415-3537. [doi: 10.1093/eurheartj/ehae177] [Medline: 39210710]
8. Gandhi S, Chen S, Hong L, et al. Effect of mobile health interventions on the secondary prevention of cardiovascular disease: systematic review and meta-analysis. Can J Cardiol 2017 Feb;33(2):219-231. [doi: 10.1016/j.cjca.2016.08.017] [Medline: 27956043]
9. Akinosun AS, Polson R, Diaz-Skeete Y, et al. Digital technology interventions for risk factor modification in patients with cardiovascular disease: systematic review and meta-analysis. JMIR Mhealth Uhealth 2021 Mar 3;9(3):e21061. [doi: 10.2196/21061] [Medline: 33656444]
10. Franssen WMA, Franssen G, Spaas J, Solmi F, Eijnde BO. Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and meta-analysis of randomised controlled trials. Int J Behav Nutr Phys Act 2020 May 11;17(1):57. [doi: 10.1186/s12966-020-00955-2] [Medline: 32393357]
11. Patterson K, Davey R, Keegan R, Freene N. Smartphone applications for physical activity and sedentary behaviour change in people with cardiovascular disease: a systematic review and meta-analysis. PLOS ONE 2021;16(10):e0258460. [doi: 10.1371/journal.pone.0258460] [Medline: 34634096]
12. Kaihara T, Intan-Goey V, Scherrenberg M, Falter M, Frederix I, Dendale P. Impact of activity trackers on secondary prevention in patients with coronary artery disease: a systematic review and meta-analysis. Eur J Prev Cardiol 2022 May 25;29(7):1047-1056. [doi: 10.1093/eurjpc/zwab146] [Medline: 34472613]
13. Ambrosetti M, Abreu A, Corrà U, et al. Secondary prevention through comprehensive cardiovascular rehabilitation: From knowledge to implementation. 2020 update. A position paper from the Secondary Prevention and Rehabilitation Section of the European Association of Preventive Cardiology. Eur J Prev Cardiol 2021 May 14;28(5):460-495. [doi: 10.1177/2047487320913379] [Medline: 33611446]
14. McDonagh ST, Dalal H, Moore S, et al. Home-based versus centre-based cardiac rehabilitation. Cochrane Database Syst Rev 2023 Oct 27;10(10):CD007130. [doi: 10.1002/14651858.CD007130.pub5] [Medline: 37888805]
15. Ramachandran HJ, Jiang Y, Tam WWS, Yeo TJ, Wang W. Effectiveness of home-based cardiac telerehabilitation as an alternative to Phase 2 cardiac rehabilitation of coronary heart disease: a systematic review and meta-analysis. Eur J Prev Cardiol 2022 May 25;29(7):1017-1043. [doi: 10.1093/eurjpc/zwab106] [Medline: 34254118]
16. Tromp J, Jindal D, Redfern J, et al. World Heart Federation roadmap for digital health in cardiology. Glob Heart 2022;17(1):61. [doi: 10.5334/gh.1141] [Medline: 36051317]
17. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
18. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, Van Spall HGC. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74. [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]

19. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
20. Niebauer J. Cardiac rehabilitation in Austria. Wien Med Wochenschr 2018 Feb;168(1-2):46-49. [doi: 10.1007/s10354-017-0607-x] [Medline: 29101493]
21. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. JMIR Mhealth Uhealth 2020 Feb 20;8(2):e15935. [doi: 10.2196/15935] [Medline: 32130167]
22. Klack L, Ziefle M, Wilkowska W, Kluge J. Telemedical versus conventional heart patient monitoring: a survey study with German physicians. Int J Technol Assess Health Care 2013 Oct;29(4):378-383. [doi: 10.1017/S026646231300041X] [Medline: 24290330]
23. Morris J, Thompson N, Wallace T, Jones M, DeRuyter F. Survey of rehabilitation clinicians in the United States: barriers and critical use-cases for mrehab adoption. In: Miesenberger K, Manduchi R, Covarrubias Rodriguez M, Peňáz P, editors. Comput Help People Spec Needs: Springer International Publishing; 2020:250-258. [doi: 10.1007/978-3-030-58805-2_30]
24. Robinson J. Likert scale. In: Maggino F, editor. Encyclopedia of Quality of Life and Well-Being Research: Springer International Publishing; 2023:3917-3918. [doi: 10.1007/978-3-031-17299-1_1654]
25. Kulnik ST. Online survey of healthcare professionals on the use of digital technology in the secondary prevention of cardiovascular disease in Austria. OSF 2024 [FREE Full text] [doi: 10.17605/OSF.IO/N9JZ6]
26. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
27. Karrer K, Glaser C, Clemens C, Bruder C. Technikaffinität erfassen - der Fragebogen TA-EG [Book in German]. In: Lichtenstein A, Stößel C, Clemens C, editors. Der Mensch Im Mittelpunkt Technischer Systeme - 8 Berliner Werkstatt Mensch-Maschine-Systeme: VDI; 2009:196-201.
28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
29. Von der Heidt A, Ammenwerth E, Bauer K, et al. HerzMobil Tirol network: rationale for and design of a collaborative heart failure disease management program in Austria. Wien Klin Wochenschr 2014 Nov;126(21-22):734-741. [doi: 10.1007/s00508-014-0665-7] [Medline: 25392254]
30. Heartfish [Website in German]. URL: https://www.heartfish.io/ [accessed 2025-01-31]
31. Mannheim I, Köttl H. Ageism and (successful) digital engagement: a proposed theoretical model. Gerontologist 2024 Sep 1;64(9):gnae078. [doi: 10.1093/geront/gnae078] [Medline: 38874215]
32. Konttila J, Siira H, Kyngäs H, et al. Healthcare professionals' competence in digitalisation: a systematic review. J Clin Nurs 2019 Mar;28(5-6):745-761. [doi: 10.1111/jocn.14710] [Medline: 30376199]
33. Miech EJ, Rattray NA, Flanagan ME, Damschroder L, Schmid AA, Damush TM. Inside help: an integrative review of champions in healthcare-related implementation. SAGE Open Med 2018;6:2050312118773261. [doi: 10.1177/2050312118773261] [Medline: 29796266]
34. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. Sports Med Open 2018 Sep 3;4(1):42. [doi: 10.1186/s40798-018-0157-9] [Medline: 30178072]
35. Stefanakis M, Batalik L, Papathanasiou J, Dipla L, Antoniou V, Pepera G. Exercise-based cardiac rehabilitation programs in the era of COVID-19: a critical review. Rev Cardiovasc Med 2021;22(4):1143-1155. [doi: 10.31083/j.rcm2204123]
36. Mäder M, Timpel P, Schönfelder T, et al. Evidence requirements of permanently listed digital health applications (DiGA) and their implementation in the German DiGA directory: an analysis. BMC Health Serv Res 2023 Apr 17;23(1):369. [doi: 10.1186/s12913-023-09287-w] [Medline: 37069592]
37. Etiwy M, Akhrass Z, Gillinov L, et al. Accuracy of wearable heart rate monitors in cardiac rehabilitation. Cardiovasc Diagn Ther 2019 Jun;9(3):262-271. [doi: 10.21037/cdt.2019.04.08] [Medline: 31275816]
38. Cohen Rodrigues TR, de Buisonjé DR, Keesman M, et al. Facilitators of and barriers to lifestyle support and ehealth solutions: interview study among health care professionals working in cardiac care. J Med Internet Res 2021 Oct 15;23(10):e25646. [doi: 10.2196/25646] [Medline: 34652280]
39. Attig C, Franke T. I track, therefore I walk – Exploring the motivational costs of wearing activity trackers in actual users. Int J Hum Comput Stud 2019 Jul;127:211-224. [doi: 10.1016/j.ijhcs.2018.04.007]
40. Zeller A, Gutenberg J, Niebauer J, Crutzen R, Kulnik ST. Patients' experiences and perspectives regarding the use of digital technology to support exercise-based cardiac rehabilitation: a qualitative interview study. Front Sports Act Living 2024;6:1371652. [doi: 10.3389/fspor.2024.1371652] [Medline: 38567184]
41. Gibson I, Kerins C, Foley L, et al. Factors Influencing Implementation of Digital Cardiac Rehabilitation: A Qualitative Analysis of Health Workers Perspectives: HRB Open Research; 2024. [doi: 10.12688/hrbopenres.13929.1]
42. Talevski J, Kulnik ST, Jessup RL, Falls R, Cvetanovska N, Beauchamp A. Use of co-design methodology in the development of cardiovascular disease secondary prevention interventions: a scoping review. Health Expect 2023 Feb;26(1):16-29. [doi: 10.1111/hex.13633] [Medline: 36366855]

43. Preusse KC, Mitzner TL, Fausset CB, Rogers WA. Older adults' acceptance of activity trackers. J Appl Gerontol 2017 Feb;36(2):127-155. [doi: 10.1177/0733464815624151] [Medline: 26753803]
44. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8]
45. Wali S, Guessi Margarido M, Shah A, et al. Expanding telemonitoring in a virtual world: a case study of the expansion of a heart failure telemonitoring program during the COVID-19 pandemic. J Med Internet Res 2021 Jan 22;23(1):e26165. [doi: 10.2196/26165] [Medline: 33444153]
46. Duscha BD, Piner LW, Patel MP, et al. Effects of a 12-week mHealth program on peak VO2 and physical activity patterns after completing cardiac rehabilitation: a randomized controlled trial. Am Heart J 2018 May;199:105-114. [doi: 10.1016/j.ahj.2018.02.001] [Medline: 29754647]
47. Lunde P, Bye A, Bergland A, Grimsmo J, Jarstad E, Nilsson BB. Long-term follow-up with a smartphone application improves exercise capacity post cardiac rehabilitation: a randomized controlled trial. Eur J Prev Cardiol 2020 Nov;27(16):1782-1792. [doi: 10.1177/2047487320905717] [Medline: 32106713]
48. Dorda W, Duftschmid G, Gerhold L, Gall W, Gambal J. Introducing the electronic health record in Austria. In: Connecting Medical Informatics and Bio-Informatics: IOS Press; 2005:119-124 URL: https://ebooks.iospress.nl/publication/10286 [accessed 2024-08-09]
49. eHealth-strategie österreich [Report in German]. : Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz; 2024.
50. Goetz G, Jeindl R, Panteli D, Busse R, Wild C. Digital health applications (DiHA): approaches to develop a reimbursement process for the statutory health insurance in Austria. Health Policy Technol 2023 Sep;12(3):ne. [doi: 10.1016/j.hlpt.2023.100780] [Medline: 37732005]
51. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144. [doi: 10.2196/34144] [Medline: 35311682]
52. Hernandez MF, Rodriguez F. Health techequity: opportunities for digital health innovations to improve equity and diversity in cardiovascular care. Curr Cardiovasc Risk Rep 2023;17(1):1-20. [doi: 10.1007/s12170-022-00711-0] [Medline: 36465151]
53. Bachner F, Bobek J, Habimana K, et al. Austria: health system review. Health Syst Transit 2018 Aug;20(3):1-254. [Medline: 30277215]


## Abbreviations

**CR:** cardiac rehabilitation
**CVD:** cardiovascular disease
**DT:** digital technology
**HCP:** health care professional
**PA:** physical activity
**WHO:** World Health Organization


*Edited by G Krstačić; submitted 16.01.25; peer-reviewed by E Amini-Salehi, L Kayser, L Allan; revised version received 22.04.25; accepted 14.05.25; published 25.06.25.*

*Please cite as:*
*Lunz L, Würth S, Kulnik ST*
*Health Care Professionals' Use of Digital Technology in the Secondary Prevention of Cardiovascular Disease in Austria: Online Survey Study*
*JMIR Cardio 2025;9:e71366*
*URL: https://cardio.jmir.org/2025/1/e71366*
*doi:10.2196/71366*

Original Paper

# Technology Readiness Level and Self-Reported Health in Recipients of an Implantable Cardioverter Defibrillator: Cross-Sectional Study

Natasha Rosenmeier[1,2*], MSc; David Busk[1,2*], MSc; Camilla Dichman[2], MScN; Kim Mechta Nielsen[2], MSc; Lars Kayser[1], MD, PhD; Mette Kirstine Wagner[2], MSc, PhD

[1]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[2]Department of Cardiology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

[*]these authors contributed equally

**Corresponding Author:**
Natasha Rosenmeier, MSc
Department of Public Health
University of Copenhagen
Øster Farimagsgade 5
Copenhagen, 1353
Denmark
Phone: 45 35334448
Email: sncb91@gmail.com

## Abstract

**Background:** Approximately 200,000 implantable cardioverter defibrillators (ICDs) are implanted annually worldwide, with around 20% of recipients experiencing significant psychological distress. Despite this, there are no ICD guidelines addressing mental health as part of rehabilitation programs, which primarily focus on educating patients about their condition and prognosis. There is a need to include elements such as emotional distress, social interactions, and the future use of technologies like apps and virtual communication in ICD rehabilitation, without increasing the burden on health care professionals.

**Objective:** This study aimed to demonstrate how data from the Readiness for Health Technology Index (READHY), combined with sociodemographic characteristics and exploratory interviews, can be used to construct profiles of recipients of an ICD, describing their ability to manage their condition, their need for support, and their digital health literacy. This aims to enhance health care professionals' understanding of different patient archetypes, serving as guidance in delivering personalized services tailored to the needs, resources, and capabilities of individual recipients of ICDs.

**Methods:** Overall, 79 recipients of an ICD participated in a survey assessing technology readiness using the READHY. The survey also collected sociodemographic data such as age, sex, and educational level. Self-reported health was measured using a Likert scale. Cluster analysis categorized participants into profiles based on their READHY scores. Correlations between READHY scores and self-reported health were examined. In addition, qualitative interviews with representatives from different readiness profiles provided deeper insights.

**Results:** Four technology readiness profiles were found: (1) profile 1 (low digital health literacy, insufficient on 5 dimensions), (2) profile 2 (sufficient on all dimensions), (3) profile 3 (consistently sufficient readiness on all dimensions), and (4) profile 4 (insufficient readiness on 9 dimensions). Participants in profile 4, characterized by the lowest readiness levels, were significantly younger ($P$=.03) and had lower self-reported health ($P$<.001) than those in profile 3. A correlation analysis revealed that higher READHY scores were associated with better self-reported health across all dimensions. Qualitative interviews highlighted differences in self-management approaches and the experience of support between profiles, emphasizing the essential role of social support toward the rehabilitation journeys of recipients of an ICD. Two patient vignettes were created based on the characteristics from the highest and lowest profiles.

**Conclusions:** Using the READHY instrument to create patient profiles demonstrates how it can be used to make health care professionals aware of specific needs within the group of recipients of an ICD.

***(JMIR Cardio 2025;9:e58219)*** doi:10.2196/58219

**KEYWORDS**

implantable cardioverter defibrillator; health literacy; self-management; ICD rehabilitation; digital health literacy; patient-reported outcome measure; self-reported; self-rated; exploratory; interview; sociodemographic; survey; cluster analysis; mixed method; cross-sectional; Denmark

## Introduction

Worldwide, approximately 200,000 implantable cardioverter defibrillators (ICDs) for primary and secondary prophylactic indications are implanted every year [1]. In Denmark, 2000 people were treated with an ICD in 2020 [2]. It is evident that implantation of an ICD with a primary prophylactic indication significantly improves the survival of patients with high-risk cardiovascular conditions who have symptomatic heart failure and a left ejection fraction below 35% [3]. Despite a significant benefit on reduction in mortality in recipients of an ICD [4] and the fact that most recipients effectively adapt to life with an ICD [5], a systematic review involving 45 studies and over 5000 recipients found that approximately 20% of recipients of an ICD experience clinically significant psychological distress [6]. Despite the acknowledged issue, there are currently no national or international ICD guidelines that specifically address the management of mental health issues as an integral component of rehabilitation. Previously, it has been proposed that rehabilitation programs should incorporate customized, hospital-based services tailored to the unique requirements and preferences of recipients of an ICD, with the aim of ensuring adequate psychological well-being and overall quality of life [5,7]. Currently, the initial rehabilitation program after discharge comprises activities aimed at enhancing understanding of the underlying disease and prognosis, as well as preparing the recipient for life with an ICD. However, there is a need to incorporate specific elements addressing the individual's unique challenges, such as emotional distress, perceived lack of support, or other person-specific concerns [8]. This necessitates the development of innovative approaches in clinical care and rehabilitation without increasing the demand for additional hours from health care professionals. A study involving individuals with chronic obstructive pulmonary disease [9] recommends incorporating both virtual and in-person components to enhance adherence [10]. To obtain the benefits of this approach, we suggest implementing similar strategies in ICD rehabilitation, as shown to be beneficial in the chronic obstructive pulmonary disease study.

When proposing the use of digital services and technology, it should be noted that approximately one-third of the older adult population in Denmark lacks a sufficient level of health literacy or digital health literacy [11]. It may be assumed that a significant number of recipients of an ICD are also challenged if expected to actively engage with digital health information. This number may even increase if the recipients are expected to participate in web-based activities in relation to a rehabilitation program. However, the challenge may be greater for recipients of an ICD than for other groups with long-term health conditions, as many recipients of an ICD are burdened by cognitive impairment as a consequence of a recent cardiac arrest, heart failure, general arteriosclerotic disease, or psychological distress [12,13]. We consider it essential, in the design of a new rehabilitation program, to address the individual needs of recipients of an ICD in relation to the heterogeneity of this group, with respect to their ability to manage their condition, their need for support, and their digital competencies. Such a redesign will enhance both the patient experience and assist in a more efficient allocation of health care professional's resources. This may involve providing virtual or even generative artificial intelligence–based services to individuals who are digitally literate and allocating in-person hours to those who require more personal contact due to social exclusion. Based on previous research involving patients with inflammatory bowel disease [14], patients with type 2 diabetes mellitus [15], and cancer survivors [16], we hypothesize that by using a patient-reported outcome dataset, such as the Readiness and Enablement Index of Health Technology (READHY) [16], alongside supplementary data on sociodemographic characteristics, it is feasible to map individuals' perceived support, self-management capabilities, and digital health literacy. This approach can facilitate the creation of patient profiles, thereby enhancing health care professionals' awareness of the diverse needs of their patients.

The READHY is a validated instrument that consists of 13 dimensions with a total of 65 items related to self-management, social support, and digital health literacy. The instrument builds on the concept of digital health literacy as the core measured with the validated eHealth Literacy Questionnaire (eHLQ; 7 dimensions), supplemented with 4 dimensions reporting on aspects of self-management from the Health Education Impact Questionnaire (heiQ) and 2 dimensions reporting on support from the Health Literacy Questionnaire (HLQ) [17-19].

The purpose of this study is to demonstrate, in the context of recipients of an ICD, how READHY data, supplemented with sociodemographical characteristics and explorative interviews, can be used to create profiles of recipients of an ICD, describing their needs, resources, and capabilities with respect to their technology readiness.

## Methods

### Study Design

The study consisted of a mixed methods, cross-sectional design in 2 parts; part one encompassed a quantitative analysis, while part two involved a qualitative inquiry. In the first part, the analysis of READHY data led to the creation of 4 profiles based on participants' self-management capabilities, perceived support levels, and digital health literacy (technology readiness). Subsequently, individuals representing high and low levels of technology readiness were invited for interviews. This approach was used to provide a voice to these profiles and to illustrate the varying perspectives within the group of recipients of an ICD.

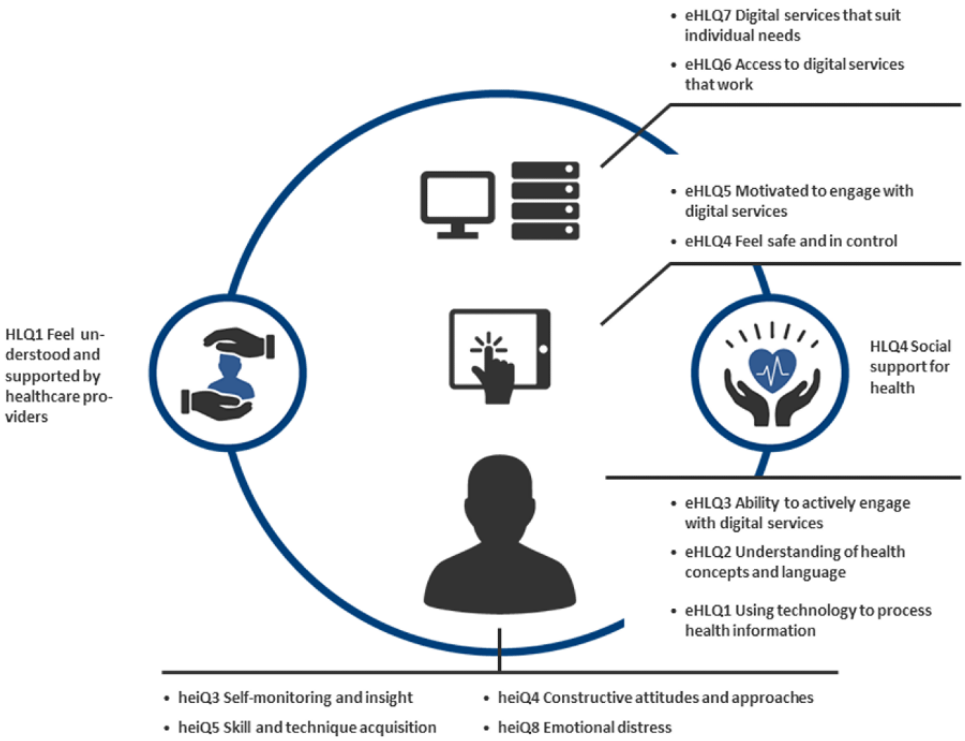
## Setting, Recruitment, and Participants

Participants included in this study were recipients of an ICD who participated in the voluntary ICD rehabilitation meeting following implantation at the Department of Cardiology at the University Hospital of Copenhagen, Rigshospitalet. The ICD rehabilitation meetings were conducted on a monthly basis, and each recipient attended only once after their device implantation. The purpose of the meeting was to address common questions about living with an ICD; provide general information and guidance about the technology behind the ICD; and explore how the treatment affects both the patient and their close relatives, including both physical and mental health issues. The meetings were facilitated only in person and by specially trained nurses, physiotherapists, and ICD technicians from the Department of Cardiology. Eligible participants were adults with primary and secondary prophylactic indications. During the research period, a total of 743 ICD devices were implanted. All patients received verbal information about the voluntary ICD rehabilitation meetings before discharge. At their first post-ICD visit, they were provided with a written invitation to the available meetings. A total of 82 (11%) patients out of 743 attended the meetings, where all completed the READHY assessment. Of these, 3 were excluded: one received a pacemaker instead of an ICD, one did not complete all of the READHY assessment, and one attended the meeting twice. The meetings were not formal hospital appointments but were offered as an additional resource for patients seeking further support and information. The inclusion took place from November 2019 to May 2022. In November 2021, a total of 6 participants, selected from a pool of 38 individuals, were invited to take part in individual semistructured interviews. In total, 3 recipients were identified from a profile of 26 individuals characterized by high levels of technology readiness, while the other 3 recipients were identified from a profile of 12 individuals with particularly low levels of technology readiness. The selection and invitation of participants was facilitated by the author, MKW, among those still in an active follow-up program at Rigshospitalet.

## Sociodemographic and Technology Readiness

A survey consisting of the READHY, sociodemographic characteristics, and self-reported health were administered at the meetings [19]) consist of between 4 and 6 items, which all have a 4-point response scale ranging from “strongly disagree” to “strongly agree.” An average score ranging from 1 (strongly disagree) to 4 (strongly agree) was calculated for each of the dimensions. The heiQ8 “emotional distress” dimension is reversed by subtracting the scores from a value of 5 for the purpose of analysis, as normally a high score would mean a high level of distress. The reversed scale now means a high level of distress has the lowest score equal to 1, so a higher score means less emotional distress as reported in the validation of the instrument [16].

**Figure 1.** The 13 dimensions of the READHY (reproduced from [16], which is published under Creative Commons Attribution 4.0 International License [20]). The 7 eHLQ dimensions describe users’ attributes; the intersection between users and technologies; and users’ experience of systems. The 4 HLQ dimensions add knowledge about the individuals’ capabilities to handle their condition and emotional response. The 2 eHLQ dimensions add knowledge about individuals’ social context (represented by the circle encompassing the individual and the individual’s attributes). eHLQ: eHealth Literacy Questionnaire; heiQ: Health Education Impact Questionnaire; HLQ: Health Literacy Questionnaire; READHY: Readiness and Enablement Index for Health Technology.




Self-rated health was assessed using a single item from the 36-item Short Form Health Survey [21]. The response options ranged from “very bad” to “very good,” graded on a scale from 1 to 5, with values of 1 to 3 indicating low self-reported health

and values of 4 to 5 indicating high self-reported health. Age was recorded in years, and sex was categorized as male or female. The response options for educational level were reported based on the International Classification of Education [22]. The 5 levels were "workers education" (eg, waiter), "skilled in craftsmanship," "short-cycle higher education," "medium-cycle higher education," and "longer education." Low educational level was categorized as scores of 1-3 and high educational level was categorized as scores of 4-5.

### Data Analysis

Data were presented as mean (SD) for continuous variables and numbers (proportions) for frequencies. Pearson product-moment correlation *r* was used to examine the correlation between self-rated health and READHY values. The degree of the correlation was defined by the *r* value, with 0.10 to 0.29 being weak, 0.30 to 0.49 being moderate, and 0.50 to 1.00 being a strong correlation [23]. Welch 2-sample *t* test (2-tailed) was used to compare READHY scores between recipients with primary and secondary prophylactic ICD indication.

### Cluster Analysis

Individuals were divided into profiles using k-means cluster analysis based on their READHY scores. The objective of the cluster analysis was to identify a profile characterized by particularly low response values across all READHY dimensions. Given the consistently low response values, this group was considered to be of particular clinical relevance for examination and comparison with profiles displaying higher response values.

Performing a k-means cluster analysis requires a prespecification of the number of clusters before the analysis can be conducted. K-means cluster analysis with 3, 4, and 5 clusters were tested in 10 iterations to determine which number of clusters had the most clinically relevant distribution. The seed value of this distribution was then saved, so that all future calculations were made from the same distribution.

Differences among the identified profiles concerning their sociodemographic characteristics and ICD indication were assessed using the Fisher exact test for categorical variables and one-way ANOVA for continuous variables. The results of the one-way ANOVA were presented with *P* values, effect size was calculated as eta-square ($\eta^2$), and Tukey multiple comparisons of means were used to assess which groups means differed significantly from each other.

Statistical calculations were performed using R (version 1.4.1717; R Core Team).

### Explanatory Interviews

This section is reported according to COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [24]. Individual semistructured interviews were conducted with 6 participants recruited as described above. All interviews were conducted in person, at a location selected by the participant (home, hospital, or university). The interviews were led by the first author, NR (female), who had no previous relationship with the participants. Each interview began with a thorough introduction to the project including the purpose of interviewing and the professional background of the interviewer. Furthermore, participants were informed that the interview was being recorded for the purpose of transcribing the conversation for further analysis. In this context, the elements of the consent form and information sheet were reviewed with the participant. Present at the interviews were the participant and the 2 first authors, NR and DB. Field notes were made during the interview by DB. The interviewer, NR, holding a master's degree in health informatics from the University of Copenhagen, is trained in conducting qualitative analyses. In addition, throughout the entire research period, the interviewer received continuous supervision from experienced researchers within the author group, LK and MKW.

A guide for the semistructured interviews was developed based on the READHY framework (Multimedia Appendix 1). The intention of the interviews was to explore the participant's perspectives on becoming a recipient of an ICD. The interview duration varied from 30 to 60 minutes, with a mean duration of 44.5 (SD 10.81) minutes. Interviews were conducted at various locations, including the hospital (n=2), the patients' homes (n=3), and at the university (n=1), accommodating the preferences of the individual participants.

Following the conclusion of each interview, a verbatim transcription was meticulously generated from the digital audio recordings. This transcription process ensured that data were accurately and comprehensively captured for subsequent analysis. The analysis of the interview data was carried out using a content analysis with an abductive approach [25]. The software package NVivo12 (Lumivero) was used. The coding was based on the READHY framework with the main categories: self-management (6 notes), social support (4 notes), and digital health literacy (4 notes). Participants have not been presented with the transcribed data nor provided feedback on the findings.

### Ethical Considerations

This study adheres to the ethical principles outlined in the Declaration of Helsinki [26]. The Danish Data Protection Agency approved the handling of data under journal P-2019-78, I-Suite 6423. Furthermore, permission to conduct the study was obtained from the heads of the Department of Cardiology at Rigshospitalet. All participants provided individual written informed consent before completing the questionnaire and participating in the interviews. Participants were informed of the voluntary nature of their participation, their right to withdraw at any time, and how their data would be used for research purposes.

According to section 14(2) of the Danish Act on Committees, health science questionnaire surveys and interview studies that do not involve human biological material do not require reporting or approval from the Danish National Centre for Ethics. Due to this exception, there were no approvals required.

All data collected were anonymized to ensure confidentiality. Personal identifiers were removed, and all data were stored securely in compliance with General Data Protection Regulation and institutional data protection regulations. The data were only

accessible to the research team, ensuring the participants' privacy was maintained.

No compensation was provided to participants for their involvement in this study. However, participants were made aware that their participation would contribute to advancing knowledge in ICD rehabilitation and the potential implementation of digital tools in the rehabilitation process.

# Results

## Overview

In total, 79 participants were included in this study. The participating recipients had a total of 29 primary and 47 secondary prophylactic indications. In 3 participants, the device indication was unknown.

## Sociodemographic Characteristics

The mean age of the 79 participants who completed the survey was 60.4 (SD 12.3) years. The distribution was 73% (56/77) male, and 63% (49/78) had a secondary prophylactic ICD indication. The participants originated from the Capital Region of Denmark and the region of Zealand, Denmark.

## Comparison of READHY Scores and Prophylactic ICD Indication

A comparison of READHY scores of those with primary and secondary prophylactic ICD indications is shown in Table 1. Lower READHY scores were observed for all 13 READHY dimensions for those with primary prophylactic indications compared to those with secondary prophylactic indications, which were significant for HQL1 ($P$=.01), HLQ4 ($P$<.001), eHLQ2 ($P$=.03), eHLQ4 ($P$<.001), and eHLQ6 ($P$=.05).

**Table 1.** Comparison of READHY[a] scores of recipients with primary and secondary prophylactic ICD[b] indication (N=76).

| READHY dimensions | *P* value | Primary prophylactic indication | Secondary prophylactic indication |
|---|---|---|---|
| heiQ[c]3: Self-monitoring and insight | .46 | 2.95 | 3.02 |
| heiQ4: Constructive Attitudes and Approaches | .09 | 3.01 | 3.14 |
| heiQ5: Skill and Technique Acquisition | .97 | 2.85 | 2.95 |
| heiQ8: Emotional Distress (reversed scale) | .98 | 2.77 | 2.95 |
| HLQ[d]1: Feeling understood and supported by healthcare providers | .01 | 3.03 | 3.23 |
| HLQ4: Social support for health | <.001 | 2.89 | 3.46 |
| eHLQ[e]1: Using technology to process health information | .69 | 2.81 | 2.99 |
| eHLQ2: Understanding of health concepts and language | .03 | 3.01 | 3.17 |
| eHLQ3: Ability to actively engage with digital services | .22 | 2.96 | 3.09 |
| eHLQ4: Feel safe and in control | <.001 | 3.13 | 3.31 |
| eHLQ5: Motivated to engage with digital services | .14 | 2.88 | 3.1 |
| eHLQ6: Access to digital services that work | .05 | 2.99 | 3.16 |
| eHLQ7: Digital services that suit individual needs | .77 | 2.79 | 2.98 |

[a]READHY: Readiness for Health Technology Index.

[b]ICD: implantable cardioverter defibrillator.

[c]heiQ: Health Education Impact Questionnaire.

[d]HLQ: Health Literacy Questionnaire.

[e]eHLQ: eHealth Literacy Questionnaire.

## READHY for Health Technology

Table 2 displays 4 health technology readiness profiles, organized in ascending order based on their average READHY scores. Profile 3 consistently exhibited sufficiency across all scales, while profile 2 was not only lower than profile 3 mostly in eHealth dimensions but also showed a sufficient level across all scales. Profile 1 showed a sufficient level on scales related to self-management and support, but insufficient levels on 5 eHealth Literacy scales except on eHLQ4 and eHLQ2. Profile 4 showed a generally insufficient level across the scales, except on HLQ1, eHLQ2, eHLQ4, and eHLQ5.

**Table 2.** Four health technology readiness profiles on the READHY[a] scale ranged from 1 (Strongly disagree) to 4 (Strongly agree; N=79). Profiles are listed from the lowest average score (left) to the highest scores (right)—highlighting the difference between each profile.

| READHY dimensions | Profiles | | | |
|---|---|---|---|---|
| | 4 (n=12) | 1 (n=9) | 2 (n=32) | 3 (n=26) |
| **Self-management, mean score** | | | | |
| hei[b]Q3 (Self-monitoring and insight) | 2.69 | 3.04 | 2.87 | 3.26 |
| heiQ4 (Constructive Attitudes and Approaches) | 2.35 | 3.16 | 2.93 | 3.65 |
| heiQ5 (Skill and Technique Acquisition) | 2.21 | 2.97 | 2.81 | 3.36 |
| heiQ8 (Emotional Distress; reversed) | 1.80 | 3.56 | 2.80 | 3.35 |
| **Support, mean score** | | | | |
| HLQ[c]1 (Feeling understood and supported by healthcare providers) | 2.77 | 3.17 | 2.97 | 3.55 |
| HLQ4 (Social support for health) | 2.58 | 3.13 | 3.19 | 3.68 |
| **eHealth literacy, mean score** | | | | |
| eHLQ[d]1 (Using technology to process health information) | 2.67 | 2.31 | 2.76 | 3.51 |
| eHLQ2 (Understanding of health concepts and language) | 2.82 | 2.84 | 2.93 | 3.58 |
| eHLQ3 (Ability to actively engage with digital services) | 2.60 | 2.35 | 2.93 | 3.67 |
| eHLQ4 (Feel safe and in control) | 3.08 | 2.87 | 3.06 | 3.73 |

[a]READHY: Readiness for Health Technology Index.

[b]heiQ: Health Education Impact Questionnaire.

[c]HLQ: Health Literacy Questionnaire.

[d]eHLQ: eHealth Literacy Questionnaire.

## Characteristics of Profiles

Differences in sociodemographic characteristics between profiles are presented in Table 3. A difference in age ($F_{3,70}$=3.1, $P$=.03, $\eta^2$=0.12) was observed. The biggest difference in age was observed between profile 4 and profile 3 ($P$=.03) and between profile 4 and profile 1 ($P$=.07). A difference in self-rated health ($F_{3,75}$=6.4, $P$=.001, $\eta^2$=0.20) was observed between the 4 profiles. The biggest difference in self-rated health was observed between profile 4 and profile 3 ($P$<.001) and between profile 3 and profile 2 ($P$=.01). No difference in sex and educational level was found. When examining for differences between the profiles with respect to ICD indication, no significant differences were found ($P$=.62). However, the percentage receiving the ICD on primary prophylactic indication in the “low-level group” was 50% (6/12) compared with the “high-level group” with only 23% (6/26). Self-rated health and level of education are measured and presented as described in the methods.

**Table 3.** Sociodemographic characteristics of participants (N=79) across profiles. Data are presented as mean (SD) for continuous variables and numbers (proportions) for frequencies.

| Characteristics | All (N=79) | Profile 1 (n=9, 11%) | Profile 2 (n=32, 40%) | Profile 3 (n=26, 33%) | Profile 4 (n=12, 15%) | *P* value |
|---|---|---|---|---|---|---|
| **Gender, n (%)** | | | | | | .45 |
| Women | 21 (27) | 1 (11) | 8 (25) | 9 (35) | 3 (25) | |
| Men | 56 (71) | 8 (89) | 24 (75) | 15 (58) | 9 (75) | |
| Unknown sex | 2 (2) | 0 (0) | 0 (0) | 2 (8) | 0 (0) | |
| **Age (years), mean (SD)** | 60.38 (12.3) | 66 (10.0) | 63 (12.7) | 58 (12.8) | 53 (7.8) | .03 |
| **Highest attained level of education, n (%)** | | | | | | .27 |
| Long education | 29 (37) | 4 (44) | 10 (31) | 12 (46) | 3 (25) | |
| Short education | 40 (51) | 4 (44) | 20 (62) | 11 (42) | 5 (42) | |
| Unknown education | 10 (13) | 1 (11) | 2 (6) | 3 (12) | 4 (33) | |
| **Self-rated health, n (%)** | | | | | | .001 |
| High self-rated health | 43 (54) | 6 (67) | 15 (47) | 20 (77) | 2 (17) | |
| Low self-rated health | 36 (46) | 3 (33) | 17 (53) | 6 (23) | 10 (83) | |
| **Prophylactic indication, n (%)** | | | | | | .62 |
| Primary | 29 (37) | 4 (44) | 13 (41) | 6 (23) | 6 (50) | |
| Secondary | 47 (60) | 5 (56) | 18 (56) | 18 (69) | 6 (50) | |
| Unknown | 3 (4) | 0 (0) | 1 (3) | 2 (8) | 0 (0) | |

## Interview Findings

To explore how differences in READHY scores related to the participants' experiences of becoming recipients of an ICD, we conducted interviews with representatives from profile 3 and profile 4. Profile 4, characterized by the lowest scores in 12 out of 13 READHY scales and lowest self-rated health, was contrasted with profile 3, which demonstrated the highest scores in all 13 scales as well as self-rated health. For the interviews, we recruited 3 participants from profile 3, here on after referred to as the "high-level group," and 3 participants from profile 4, here on after referred to as the "low-level group." These interviews revealed significant differences in how individuals from these groups were able to manage their condition, perceived the support they received, and approached digital proficiency.

## Self-Management

All participants engaged in self-management practices addressing their physical and mental well-being. However, there was a distinction in how self-management was interpreted within the "high-level group" compared to the "low-level group." Participants belonging to the "high-level group" described their pre-ICD implantation lifestyle as characterized by daily physical exertion, which they expressed a strong desire to sustain. For instance, P3 stated:

> *I used to bike to work throughout the year, covering approximately 10 kilometers each way. I engaged in workouts at least twice a week and participated in a weekly spinning class. Exercise, to me, equates to an enhanced quality of life, both presently and prior to my illness. At present, I attend one or two spinning classes weekly, which I prefer not to disclose to my doctors, as they disapprove.*

In contrast, no one in the "low-level group" used physical activity as a means to preserve their health.

Participants belonging to the "low-level group" approached self-management in a distinct manner, which primarily involved adhering to medical advice regarding medication adherence and health care appointments, particularly evident when asked about their self-care practices. For example, P2 and P5 articulated:

> *After doctors' appointments I am more sensitive and attentive to my body. Naturally, the plan is to initiate lifestyle changes, which I have gradually commenced." And "It seems like that's all I'm engaged in - devoting my time to managing my health. I visit the hospital constantly, and I mean incessantly. Furthermore, I was enrolled in a heart rehabilitation program last year.*

For individuals within the "low-level group," a recurring subject was found, wherein the participants lived with constant awareness and apprehension regarding their condition. For instance, when asked, "During your daily routine, when do you find yourself contemplating your ICD?" P1 articulated "Constantly! It occupies my thoughts incessantly." P2 concurred, stating:

> *I think about it every time I shower, change my clothing, and when I retire for the night; those are the moments when it preoccupies my mind the most. Additionally, I grapple with mental concerns such as whether it would effectively function in the event of an unforeseen circumstance.*

Similarly, P5 shared, “All the time! I am in a constant state of unease.”

When the same question was asked to participants belonging to the “high-level group,” the responses conveyed a sense of calm and trusting emotional state. As exemplified by P4 and P6:

> *My perspective has been somewhat matter of fact; I needed to have this device implanted, and that is simply the way it is. Beyond that, I have not dwelled on it extensively.* [P4]

> *After a full day at work, I may experience some soreness, but it reminds me of how reassuring it is to have it watching over me.* [P6]

## Support

### *Social Support*

In the management of their ICD, participants who felt a lack of social support from family and friends during the rehabilitation process have heightened emotional distress, necessitating additional support from health care professionals. Without substantial social support from family and friends, the perception of support from health care professionals during their hospitalization and rehabilitation process became crucial. A lack of social support affected the participant’s ability to place trust in the ICD technology and their capacity to adapt calmly to life with an ICD.

The significance of having access to supportive relatives or spouses was emphasized by the contrast in how the 2 groups used and derived comfort from sharing their concerns with close family members. The “high-level group” experienced tremendous comfort in doing so, whereas the “low-level group” tended to conceal their feelings and kept their worries to themselves. For instance, P4 remarked:

> *Discussing things with my family and my wife, who was present at the time of my cardiac arrest, and having those conversations with people who asked about my experiences, has actually proven more beneficial than speaking with the psychologist.*

This contrasted with the experiences of recipients in the “low-level group,” who perceived their condition as more burdensome for their families than as a source of support. P2 explained:

> *You may want to confide in your family, but not be completely honest about how frightened you have been and still are about the future. It's a delicate topic. My family was deeply shaken, and they may not wish to revisit it.*

Similar sentiments were expressed by P5:

> *My children are 22 and 23 years old, but they have been extremely anxious. Being a single mom and trying to stay strong for them is challenging. Yet, they want me to share my feelings. It's just very tough at times.*

### *Professional Support*

Participants who lived alone exhibited a greater demand for support and information from health care professionals when compared with participants living with a spouse. Those living alone consistently expressed dissatisfaction with the support provided by health care professionals and commonly expressed high levels of emotional distress, as well as a lack of information, support, and therapeutic options. P1 felt that his needs were overlooked and emphasized the need for more information about his condition, stating:

> *When you get admitted here, you receive absolutely no information. None. That is a flaw. I was operated on at 2 a.m., and by 9 a.m., I was approached by a professor and a nurse who wanted to recruit me for a study. That was bewildering. After surgery, your mind is in turmoil, and here they are asking me to participate in a study.*

In addition, another participant who was living alone, P5, expressed dissatisfaction with the lack of fulfillment and comprehension of her needs during her hospitalization, particularly concerning the therapy options offered after surgery. She stated:

> *During my hospitalization, I attended a few sessions with their psychologist, but it didn't resonate with me at the time. They advised me to go for forest walks and visit the library to socialize. That wasn't what I needed.*

In contrast, all participants living with a partner consistently reported the support provided by health care professionals as highly satisfactory. P4 stated:

> *I felt safe from the moment I woke up in the hospital and throughout my entire stay. I have been extremely pleased with the care and treatment I received here.*

P6 similarly expressed positive impressions, saying:

> *I wish I could write an article about it; it felt like a five-star hotel. They treated me like royalty, providing me with detailed information, time, and care. We were deeply impressed by the dedication and attention they gave us.*

### *Digital Health Literacy*

Participants from both the “high-level” and “low-level” groups expressed a consistent readiness and ability to engage with digital health care services and use various technological tools as part of their recovery process. They shared a common inclination for monitoring their health data, seeking health information online, and accessing personal health records through digital platforms. There was no noticeable difference in motivation for digital rehabilitation between the 2 groups, potentially due to their recruitment from a rehabilitation program rather than during hospitalization. Moreover, both groups displayed similar engagement with other health-related technologies, such as smartwatches and pulse oximeters, indicating their willingness to embrace technology for a digitalized rehabilitation experience tailored to their needs.

A participant belonging to the "low-level group," P5, detailed her utilization of various technologies for managing her condition:

> *I have been using my Apple Watch since I received my first pacemaker. Sometimes, I would feel unwell and worry about my pulse being too low. Tracking it on my watch gives me peace of mind. Additionally, I regularly log in to my online electronic health record to stay informed about any updates. The more information I acquire, the more at ease I feel.*

Similarly, P4 belonging to the "high-level group" expressed:

> *I purchased an actual pulse oximeter when my condition first arose. I told my wife that I needed one. I have an imperative need to comprehend what is transpiring.*

### *ICD Indication*

One distinguishing characteristic of recipients within the "low-level group" was their lack of trust in the ICD technology and the high levels of emotional distress they experienced living with an ICD. It is noteworthy that the 3 recipients belonging to the "low-level group" had previously been diagnosed with heart-related conditions before receiving the ICD, which contrasts with the participants belonging to the "high-level group" who had no such previous diagnoses. The recipients with an ICD who have primary prophylactic indication consistently exhibit notably low READHY scores, especially in the domain of social support, when compared to recipients with secondary prophylactic indication. Interviews show that the overall health status of the recipient before ICD placement is an essential determinant influencing the patient's ability to manage the condition. Importantly, the interviewer had no previous knowledge of which group the interviewed participants belonged to.

### *Patient Vignettes*

Based on data presented in Tables 2 and 3 and the qualitative interviews, we have created 2 patient vignettes, which are presented below. These demonstrate how the text vignettes can make the profiles more vivid for health care professionals.

#### Vignette for the Low-Level Group

This is a male individual aged 53 years with low physical activity levels and low self-rated health, diagnosed with other comorbidities before ICD implantation. The patient is unmarried, lives alone, has a limited social network, and experiences significant emotional distress due to his condition on a daily basis. He uses health technologies and actively seeks information about his condition online. The "low-level group" of patient requires a high level of support from health care professionals during hospitalization and through their rehabilitation process.

#### Vignette for the High-Level Group

This is a male individual aged 58 years with a high level of physical activity and high self-rated health, who maintains good health and has no comorbidities before his ICD implantation. The patient cohabits with a partner and has an extensive social network. He maintains a positive attitude toward his condition and incorporates health technologies into his daily routine.

## Discussion

### Principal Findings

The purpose of this study was to demonstrate how profiles and patient vignettes can be developed using the READHY instrument to make health care professionals aware of differences in patient's needs, resources, and capabilities in relation to their health technology readiness, including their emotional state. Using cluster analysis, 4 clinically relevant profiles were developed. The most distinct profiles we found were profile 3, characterized by highly sufficient READHY scores across all dimensions, and profile 4, characterized by 9 insufficient READHY scores (below 2.7), displaying only slight sufficiency within digital literacy. Sociodemographic characteristics, age, and self-reported health differed among the profiles, with the youngest patients having the lowest READHY scores. No significant differences were found in sex, level of education, or ICD indication. This underpins the need other than these classical characteristics to inform the health care professionals to understand their patients. The interviews provided valuable insights into the perspectives of the profiles, emphasizing the crucial role of social support, particularly for those living alone, who required more professional support. These insights were particularly relevant with regard to emotional distress and perceived support levels from family and health care professionals.

Individuals with no or a short history of poor health conditions tended to adapt more positively to life post-ICD implantation, compared with those with a longer history of poor health conditions. This suggests that it may be significant to take the patient's previous and current status of health into consideration in the treatment of them. Interestingly, interviewees belonging to both the low and high-level groups embraced technology to a high extent, signifying that in recipients of an ICD, physical health is not related to the usage of technology.

### Profile Characteristics

#### *Age and Self-Rated Health*

We found significant differences in age and self-reported health among the recipients of an ICD in different profiles, but no significant difference in sex, educational level, or ICD indication. Profile 4, which represents individuals with the lowest READHY scores, is comprised of individuals who are, on average, 13 years younger than those in the oldest profile. This contrasts with previous research, where older adults tended to have poorer health outcomes [15]. The youngest patients had the lowest scores in self-rated health, indicating that age alone may not be a strong predictor of ICD-related health outcomes. This suggests the importance of considering other factors such as other long-term health conditions and self-rated health status when assessing patient needs, resources, and capabilities, rather than age.

#### *Social Support*

In alignment with previous findings [15], our interview data show that emotional and social support from a partner or spouse plays a role in addressing emotional concerns after ICD placement. The participants living with a spouse reported an

exceptionally high level of received care from health care professionals and had little need to seek additional support. Conversely, participants living alone expressed feelings of abandonment, lack of information, and insufficient care from health care professionals.

The impact of social support on mental well-being is further evident in the difference in emotional concerns between the "high-level" and "low-level" groups. The "high-level group" expressed trust in their ICD and had fewer daily worries about their condition, whereas all participants in the "low-level group" reported doubts about their ICD's effectiveness and ongoing concerns about their future health. Therefore, the presence or absence of social support in the form of a spouse or near family is a crucial factor to consider when identifying patients who may require additional support and tailored rehabilitation services.

### *Digital Health Literacy*

The recipients of an ICD had relatively high levels of digital health literacy scores in both the "low-level" and "high-level" groups compared to patients with inflammatory bowel disease [14]. The sufficiency of digital health literacy was further confirmed during interviews, where all participants reported regular use of digital health tools in their daily lives. This contrasts with previous research, which suggests limited technology engagement among individuals with chronic illnesses [14]. In our study, recipients of an ICD from various profiles actively embraced technology for health monitoring; sought health-related information online; and used devices such as smartwatches, fitness trackers, and advanced pulse oximeters, regardless of their profile. This collective engagement suggests an opportunity among recipients of an ICD to adopt new digital services and technology.

Our interviews involved individuals from profiles 4 and 3. Profiles 4 and 3 were selected due to having the overall lowest and highest READHY scores, respectively, but it should be noticed that the lowest levels of digital health literacy were found in profile 1.

The characteristics of participants belonging to profiles 1 and 2 should also be considered when planning rehabilitation. Identifying individuals within these intermediate profiles is essential, as they may also exhibit low values in specific dimensions. Profile 1 had a sufficient level within the areas of self-management and social support but was found with lower levels in digital health literacy compared with the other profiles. The introduction of digital technologies may pose a barrier for this group, as they do not possess the same high levels of digital literacy as the other groups. In essence, while they excel in traditional health-related knowledge, they may struggle when it comes to using digital health tools and resources. This group should be approached recognizing their nondigital competence and with a careful introduction of digital solutions.

Profile 2 was the largest group, characterized by having sufficient levels on all scales. Despite having lower levels than those in profile 3, they are considered capable of actively participating in their rehabilitation including complementary digital services and technologies. The key here is to recognize individuals who are less capable than those in profile 4 but still require increased assistance and rehabilitation services, especially within the self-management area.

Due to the fact that recipients of an ICD can be clustered into diverse patient profiles where some have low digital literacy, we advocate retaining the in-person ICD rehabilitation meeting as an available option, particularly for individuals belonging to profiles 1 and 4. This group may benefit from additional support, counseling, and information throughout their recovery process, ensuring a more comprehensive and personalized approach to their care. The interviews indicated that all individuals, regardless of which of the 2 profiles they belonged to, regularly used digital services and found them to be comfortable and reassuring. This suggests that most recipients of an ICD, including those with lower levels of digital health literacy, can benefit from the enhanced integration of technology into the ICD rehabilitation program. Using the READHY instrument to identify profiles and their associated individuals will serve as a valuable tool in tailoring future ICD treatments to meet individual needs.

### *ICD Indication*

Regarding the differences in prophylactic indication, it is important to recognize that the current treatment pathways vary based on the indication. Patients undergoing secondary ICD placement, often due to acute conditions like cardiac arrest, experience a more prolonged hospital stay compared with those undergoing planned, elective, primary ICD placement. Conducting a study that combines both primary and secondary indications for ICD placement involves including a group of patients who have not undergone the exact same treatment process. Despite this, our qualitative analysis remained impartial, as all interviewed participants underwent secondary ICD placement, ensuring a one-to-one basis for comparison.

Recipients with primary ICD indications had lower, but sufficient, levels of all 13 READHY scales compared with those with secondary indications. This was significant in relation to support from both professionals (HLQ1) and relatives or peers (HLQ4); it was also significant in relation to the 3 digital health scales concerning having access to digital services for those who need them (eHLQ6), trusting how their data are handled (eHLQ4), and understanding the health language (eHLQ2). The higher READHY scores from recipients with a secondary indication for ICD placement could be due to their prolonged hospitalization, which gave them more extensive interaction with health care professionals. Another explanation could be that this group has not experienced a prolonged history of poor health, resulting in fewer interactions with the health care sector and potentially fostering a more optimistic outlook.

### *Patient Vignettes*

A way to make the profiles more present and recognizable by health care professionals is to create vignettes that describe a particular average person belonging to a specific profile.

The vignettes offer insights into the unique needs, challenges, and behaviors of individuals within the "low-level" and "high-level" groups of this study. By delving into the details of these vignettes, we aim to provide a deeper understanding of

how various factors, including health status, social support, and lifestyle, influence the experiences of recipients of an ICD. The vignettes serve as representative examples with the purpose of assisting health care professionals in identifying patient characteristics, ultimately enabling the delivery of more tailored support and care to the population of recipients of an ICD. It remains to be tested in a clinical setting to what extent these vignettes can help the health care professionals in their everyday work.

### Strengths and Limitations

A strength of the study lies in its foundation on an established model previously used in patients with other chronic conditions. The data help translate the understanding of health technology readiness into a new clinical area, providing a fresh perspective for health care professionals in cardiology. This enables them to better meet patients' needs while considering their resources and capabilities in a digital context, including mental and social aspects.

However, a limitation of this study is the absence of interviews with individuals from profile 1, which is characterized by the lowest level of digital health literacy, particularly in scales eHLQ1, eHLQ3, eHLQ5, and eHLQ7. Including interviews from this group could have yielded valuable insights into the factors contributing to their low digital competence. By not doing so, the depth and comprehensiveness of the data were somewhat limited.

In addition to the above, another potential limitation is the relatively low number of participants, which may introduce a risk of bias, as only those with a high level of self-management ability may have participated. This could also increase the risk of a type 2 error, potentially overlooking differences between profiles in sociodemographic characteristics and self-reported health.

Furthermore, the survey sampling took place over a period of 2 years and 7 months, during which the COVID-19 pandemic occurred, limiting the number of participants that could be included. A multicenter study would have been necessary to achieve a larger sample size within this timeframe. Nevertheless, despite this limitation, the data still contribute significantly to our understanding of recipients of an ICD and the dynamics of their competencies.

Finally, a limitation in interpreting the differences between primary and secondary indications for ICD placement is worth noting. Some individuals in the secondary group may have had preexisting heart conditions, making them more similar to patients in the primary group. Unfortunately, this factor was not accounted for in the study design, as the health care professionals involved no longer had responsibility for these patients. Although differences in READHY scales and self-rated health between the groups suggest this may have been a minor issue, future studies should emphasize assessing preexisting heart conditions and the need for cardiac resynchronization therapy.

### Conclusion

The profiles developed in this study offer a practical tool to translate complex data into a more accessible format, enabling health care professionals to identify individuals who require additional support and those who may benefit from increased online contact. These profiles can be transformed into patient vignettes, presented in a concise text format, which help clinicians recognize specific needs related to self-management, digital health literacy, and experienced support in the context of ICD rehabilitation.

For example, profile 3 demonstrated high readiness scores across all dimensions, indicating strong self-management capabilities and a potential for greater engagement with digital health tools. In contrast, profile 4 had low scores across multiple areas, representing individuals with significant challenges in managing their condition and engaging in a rehabilitation process. These profiles highlight the spectrum of readiness and the need for tailored interventions.

It is equally important to acknowledge intermediate profiles, such as profiles 1 and 2, which exhibit unique needs that demand tailored rehabilitation approaches, particularly in the context of digital health literacy. By understanding the diversity within this population and considering the impact of sociodemographic factors, health status, and social support, health care professionals can provide more personalized and effective care to recipients of an ICD in the future.

### Acknowledgments

The authors thank the recipients of an implantable cardioverter defibrillator (ICD) who participated and the ICD Team—Ditte Petersen, Mette Lund, and Emanuella Naumova—for helping with the collection of data. This project received no funding.

### Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 14 KB - cardio_v9i1e58219_app1.docx ]

### References

1. Nisam S, Reddy S. The story of … a lead. Europace 2015 May;17(5):677-688. [doi: 10.1093/europace/euu391] [Medline: 25755290]
2. Sommer A, Witt CT, Diederichsen S. Implanterbar cardioverter defibrillator (ICD) [Article in Danish]. Dansk Cardiologisk Selskab. 2024 Apr 9. URL: https://nbv.cardio.dk/icd [accessed 2023-11-02]
3. Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB, et al. 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: executive summary: a report of the American college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society. J Am Coll Cardiol 2018;72(14):1677-1749 [FREE Full text] [doi: 10.1016/j.jacc.2017.10.053] [Medline: 29097294]
4. Alhakak A, Østergaard L, Butt JH, Vinther M, Philbert BT, Jacobsen PK, et al. Cause-specific death and risk factors of 1-year mortality after implantable cardioverter-defibrillator implantation: a nationwide study. Eur Heart J Qual Care Clin Outcomes 2022;8(1):39-49. [doi: 10.1093/ehjqcco/qcaa074] [Medline: 32956442]
5. Dunbar SB, Dougherty CM, Sears SF, Carroll DL, Goldstein NE, Mark DB, et al. Educational and psychological interventions to improve outcomes for recipients of implantable cardioverter defibrillators and their families: a scientific statement from the American heart association. Circulation 2012;126(17):2146-2172. [doi: 10.1161/CIR.0b013e31825d59fd] [Medline: 23008437]
6. Magyar-Russell G, Thombs BD, Cai JX, Baveja T, Kuhl EA, Singh PP, et al. The prevalence of anxiety and depression in adults with implantable cardioverter defibrillators: a systematic review. J Psychosom Res 2011;71(4):223-231. [doi: 10.1016/j.jpsychores.2011.02.014] [Medline: 21911099]
7. Berg SK, Pedersen PU, Zwisler A, Winkel P, Gluud C, Pedersen BD, et al. Comprehensive cardiac rehabilitation improves outcome for patients with implantable cardioverter defibrillator. Findings from the COPE-ICD randomised clinical trial. Eur J Cardiovasc Nurs 2015;14(1):34-44. [doi: 10.1177/1474515114521920] [Medline: 24504872]
8. Risom SS. National klinisk retningslinje for rehabilitering til patienter med atrieflimren, atrieflagren, patienter med endokarditis og patienter behandlet med en implanterbar cardioverter defibrillator (ICD) [Article in Danish]. Danske Fysioterapeuter. URL: https://www.fysio.dk/globalassets/documents/fafo/kliniske-retningslinjer/hjerte-lunge/nkr_rehabilitering_arterieflimren_endokarditis_icd.pdf [accessed 2024-12-18]
9. Palshof MK, Jeppesen FKH, Thuesen AD, Holm CS, Brøndum E, Kayser L. Comparison of the level of eHealth literacy between patients with COPD and registered nurses with interest in pulmonary diseases. Computer Methods and Programs in Biomedicine Update 2023;4:100121. [doi: 10.1016/j.cmpbup.2023.100121]
10. Krag T, Jørgensen EH, Phanareth K, Kayser L. Experiences with in-person and virtual health care services for people with chronic obstructive pulmonary disease:qualitative study. JMIR Rehabil Assist Technol 2023;10:e43237 [FREE Full text] [doi: 10.2196/43237] [Medline: 37578832]
11. Poulsen HS, Eiriksson SD, Christiansen ASJ, Wingstrand A. Sundhedsprofil 2021 for Region Sjælland og kommuner: »Hvordan har du det?«. Sorø, Denmark: Region Sjælland, Data og udviklingsstøtte; 2022:12.
12. Hallas CN, Burke JL, White DG, Connelly DT. A prospective 1-year study of changes in neuropsychological functioning after implantable cardioverter-defibrillator surgery. Circ Arrhythm Electrophysiol 2010;3(2):170-177. [doi: 10.1161/CIRCEP.109.909580] [Medline: 20197542]
13. Kramer DB, Habtemariam D, Adjei-Poku Y, Samuel M, Engorn D, Reynolds MR, et al. The ecisions, nterventions, and oals in implatable cardioverter-defbrillator herap (DIGNITY) pilot study. J Am Heart Assoc 2017;6(9):e006881 [FREE Full text] [doi: 10.1161/JAHA.117.006881] [Medline: 28939708]
14. Nielsen AS, Hanna L, Larsen BF, Appel CW, Osborne RH, Kayser L. Readiness, acceptance and use of digital patient reported outcome in an outpatient clinic. Health Informatics J 2022;28(2):14604582221106000 [FREE Full text] [doi: 10.1177/14604582221106000] [Medline: 35658693]
15. Thorsen IK, Rossen S, Glümer C, Midtgaard J, Ried-Larsen M, Kayser L. Health technology readiness profiles among danish individuals with type 2 diabetes: cross-sectional study. J Med Internet Res 2020;22(9):e21195 [FREE Full text] [doi: 10.2196/21195] [Medline: 32930669]
16. Kayser L, Rossen S, Karnoe A, Elsworth G, Vibe-Petersen J, Christensen JF, et al. Development of the multidimensional Readiness and Enablement Index for Health Technology (READHY) tool to measure individuals' health technology readiness: initial testing in a cancer rehabilitation setting. J Med Internet Res 2019;21(2):e10377 [FREE Full text] [doi: 10.2196/10377] [Medline: 30747717]
17. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
18. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013;13:658 [FREE Full text] [doi: 10.1186/1471-2458-13-658] [Medline: 23855504]
19. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]

20. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2025-01-30]
21. Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30(6):473-483. [Medline: 1593914]
22. International Standard Classification of Education (ISCED). UNESCO Institute for Statistics. 2017. URL: https://uis.unesco.org/en/topic/international-standard-classification-education-isced [accessed 2023-11-01]
23. Pearson's correlation coefficient. Statistics Solution. URL: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/pearsons-correlation-coefficient/ [accessed 2024-01-13]
24. Booth A, Hannes K, Harden A, Noyes JP, Harris J, Tong A. COREQ (Consolidated Criteria for Reporting Qualitative Studies). In: Guidelines for Reporting Health Research: A User's Manual. Hoboken, NJ: John Wiley & Sons, Ltd; 2014:214-226.
25. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
26. WMA Declaration of Helsinki - ethical principles for medical research involving human subjects. The World Medical Association. URL: https://tinyurl.com/3vt2ytcv [accessed 2023-11-01]

## Abbreviations

**COREQ:** Consolidated Criteria for Reporting Qualitative Research
**eHLQ:** eHealth Literacy Questionnaire
**heiQ:** Health Education Impact Questionnaire
**HLQ:** Health Literacy Questionnaire
**ICD:** implantable cardioverter defibrillator
**READHY:** Readiness for Health Technology Index

*Edited by KC Wong; submitted 09.03.24; peer-reviewed by Z Geng, T Annfeldt, C Baxter; comments to author 26.06.24; revised version received 29.09.24; accepted 04.11.24; published 06.02.25.*

*Please cite as:*
*Rosenmeier N, Busk D, Dichman C, Nielsen KM, Kayser L, Wagner MK*
*Technology Readiness Level and Self-Reported Health in Recipients of an Implantable Cardioverter Defibrillator: Cross-Sectional Study*
*JMIR Cardio 2025;9:e58219*
*URL: https://cardio.jmir.org/2025/1/e58219*
*doi:10.2196/58219*
*PMID:*

# Wearable Electrocardiogram Technology: Help or Hindrance to the Modern Doctor?

Samuel Smith[1,2], MBBS(Hons), MPH; Shalisa Maisrikrod[2,3], MBBS, MPH

[1]Department of Intensive Care Medicine, Royal Brisbane and Women's Hospital, Butterfield Street, Brisbane, Australia

[2]Faculty of Medicine, University of Queensland, Brisbane, Australia

[3]Department of Internal Medicine and Aged Care, Royal Brisbane and Women's Hospital, Brisbane, Australia

**Corresponding Author:**
Samuel Smith, MBBS(Hons), MPH
Department of Intensive Care Medicine, Royal Brisbane and Women's Hospital, Butterfield Street, Brisbane, Australia

## Abstract

Electrocardiography is an essential tool in the arsenal of medical professionals, Traditionally, patients have been required to meet health care practitioners in person to have an electrocardiogram (ECG) recorded and interpreted. This may result in paroxysmal arrhythmias being missed, as well as decreased patient convenience, and thus reduced uptake. The advent of wearable ECG devices built into consumer smartwatches has allowed unparalleled access to ECG monitoring for patients. Not only are these modern devices more portable than traditional Holter monitors, but with the addition of artificial intelligence (AI)-led rhythm interpretation, diagnostic accuracy is improved greatly when compared with conventional ECG-machine interpretation. The improved wearability may also translate into increased rates of detected arrhythmias. Despite the many positives, wearable ECG technology brings with it its own challenges. Diagnostic accuracy, managing patient expectations and limitations, and incorporating home ECG monitoring into clinical guidelines have all arisen as challenges for the modern clinician. Decentralized monitoring and patient alerts to supposed arrhythmias have the potential to increase patient anxiety and health care visitations (and therefore costs). To better obtain meaningful data from these devices, provide optimal patient care, and provide meaningful explanations to patients, providers need to understand the basic sciences underpinning these devices, how these relate to the surface ECG, and the implications in diagnostic accuracy. This review article examines the underlying physiological principles of electrocardiography, as well as examines how wearable ECGs have changed the clinical landscape today, where their limitations lie, and what clinicians can expect in the future with their increasing use.

***(JMIR Cardio 2025;9:e62719)*** doi:10.2196/62719

**KEYWORDS**

mobile applications; electrocardiogram; wearable monitoring; app; wearable; electrocardiograph; ECG; electrocardiography; mobile app; tool; ischemic; arrhythmia; wearable ECG; doctor; smartwatch; atrial fibrillation

## Introduction

The electrocardiogram (ECG) is one of the most commonly obtained test results in medical practice [1,2]. By measuring the electrical activity of the heart, an ECG can indicate cardiac arrhythmias and structural defects, respiratory disease, electrolyte disturbances, and even noncardiac events such as subarachnoid hemorrhage [1]. Traditional 12-lead ECGs are obtained by placing 10 adhesive electrodes on a patient, recording 10 seconds of electrical activity, and this snapshot is recorded for interpretation [3]. With the modern explosion of portable digital technology, a single lead ECG can now be performed without adhesive electrodes on a patient, using their own smart device, and these digital ECGs can be sent across vast distances for real-time clinician interpretation anywhere, at any time [3]. Whilst early, studies have suggested that the positive predictive value for arrhythmias such as atrial fibrillation (AF) may lie between 84% and 97% [4,5]. With a range of popular wearable technologies incorporating this feature, more number of patients with low cardiac risk have continuous ECG monitoring than ever before. This, plus the increasing role of deep learning and artificial intelligence (AI) in ECG interpretation, have implications for medical practitioners. More patients will be presenting with possibly abnormal ECGs recorded by their home devices, with associated anxiety and health care use already reported [6]. It is up to physicians have a thorough understanding of the basic sciences underpinning ECG acquisition in order to provide ECG interpretation and explain how these new devices work. This article will review the fundamentals of the ECG before examining the potential impacts of the digital age on electrocardiography for the modern doctor.

### History of the ECG

This history of the ECG is really the history of electrophysiology, which can be traced back to Galvani's [7] experimentation in the 18th century on the role of electricity in

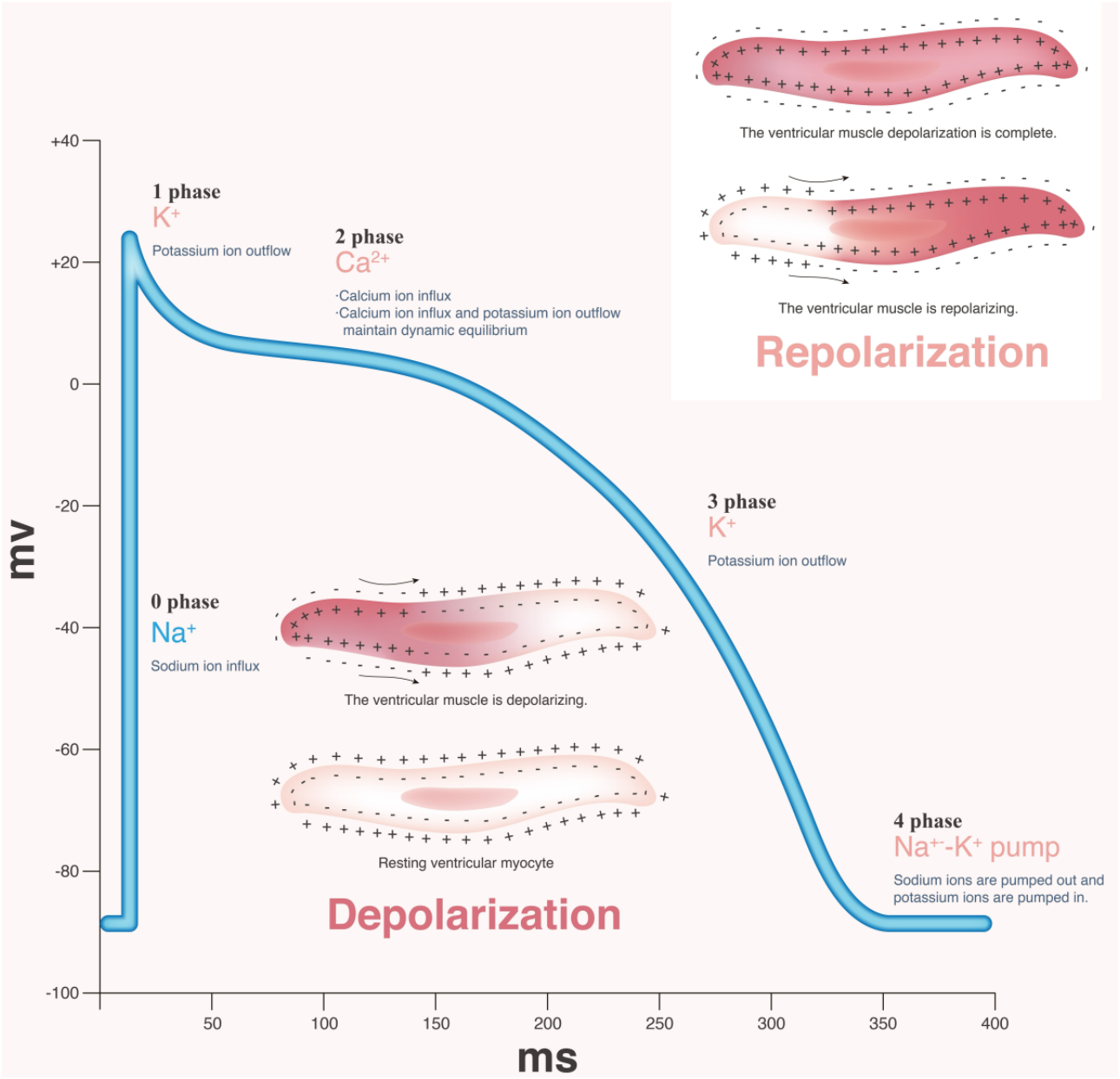
the frog nervous system. More researchers followed him, and in 1902, Einthoven broke new ground by accurately recording the electrical activity of the heart using his string galvanometer [8,9]. The string galvanometer was not without its drawbacks; it required the patient to place their hands and 1 foot into a saltwater solution, 5 assistants to operate, and weighed over 300 kilograms [10].

Thankfully, modern ECG machines have evolved, and now require only 10 small electrodes to be placed on the patient to obtain an almost complete view of the heart. Despite this, the basic principles underpinning ECG acquisition and interpretation remain unchanged since its 1902 inception, an understanding of cardiac anatomy and physiology, and physics.

## The ECG: Underlying Physiological Fundamentals

Cardiomyocytes have a positive charge on their outer membrane that result from the intra- and extracellular distribution of ions. At rest, potassium ($K^+$) ions are at a high concentration intracellularly whilst sodium ($Na^+$), calcium ($Ca^{2+}$), and chloride ($Cl^-$) have a higher concentration outside of the cell [11]. The balance of ion flow (predominantly by the outward diffusion of $K^+$ owing to membrane permeability) results in a resting membrane potential (RMP) of around −90mV [11]. Pacemaker cardiomyocytes have no stable RMP; instead, there is a constantly slowly increasing membrane potential mediated by the slow $Na^+$ "funny current" ($I_f$) [11]. Contractile myocytes are depolarized after pacemaker cells depolarize, thereby opening $I_f$ T and L-type $Ca^{2+}$ channels. Fast-$Na^+$ channels then open and allow an influx of positive $Na^+$ ions, depolarizing the cell to about +20mV and opening slow L-type $Ca^2$ channels. Once these channels close, active transports for sodium and calcium begin removing these ions to restore ionic equilibrium and a potassium rectifier channel will open, allowing $K^+$ ions to leave the cell again, repolarizing the cell (Figure 1) [12,13].

**Figure 1.** Cardiac depolarization: myocyte cardiac action potential showing ion flux across the membrane and resultant changes in the resting membrane potential and depolarization wavefront.




As each cell's membrane becomes positively charged during depolarization, they propagate their action potentials to other nearby cells, and so on. In each wavefront of depolarization, there will be positive and negative ends, which result in a moving electrical dipole [14].

A moving electrical dipole creates an electrical current. By virtue of the body's ability to act as a volume conductor, the current field created by the flow of electricity (caused by cardiac depolarization) is conducted to the thoracic cavity, and from there, the surface of the body [2,14]. This current flow is thus detectable as an electrical field on the skin by surface electrodes. The 2 electrodes act as voltmeters at their respective points and measure the potential difference between them, with the "view" between the positive and negative electrode known as a lead.For example, Lead I represents the potential difference between voltages measured at the right arm (RA; negative electrode) and left arm (LA; positive electrode) [15]. As an electric field moves toward the left arm (positive electrode), a positive potential difference (or voltage) is recorded, which would be reported as an upstroke in the ECG trace [14].

It is important to remember that there are many thousands of myocardial fibers, each with its own electrical wavefront. Surface electrodes will not be able to distinguish the electrical field generated by each wavefront, and so, the electrical field detectable on the surface of the chest wall is determined by the vectoral sum of the electromotive field strength of all active components of the myocardium [2]. It is this overall vector sum (or cardiac dipole) that is represented by the ECG trace. Having multiple leads allows simultaneous recording of the same current flow in many different views. Traditionally, a 12-lead view is used in clinical electrocardiography. This includes Einthoven's original 3-lead view, as well as 3 augmented leads (which are unipolar with a neutral central terminal) and 6 precordial leads (whose leads lie in a transverse plane) [15]. This requires the

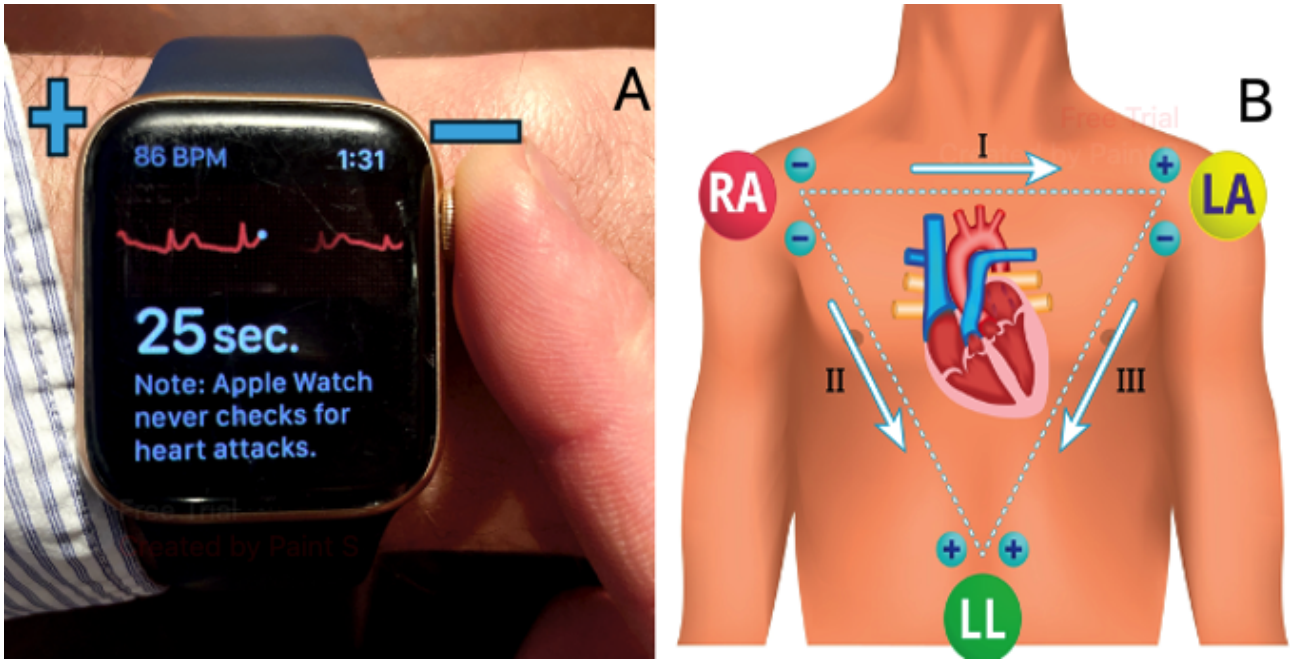
placement of 10 separate electrodes to create an electrical window for each lead [2].

## A Modern Take

Recently, breakthroughs in both the hardware and software of mobile devices have drastically changed the paradigm of ambulatory ECG monitoring, allowing ECG monitoring using wearable devices and the immediate analysis of ECGs using AI. Mobile devices are almost ubiquitous in modern society and are used daily by 2-3 billion people [16]. In a society where patients are eager for more involvement in their health and have a smartphone at their fingertips, it should come as little surprise that technology for home health monitoring has developed at a rapid pace. The wearable ECG device is an example of this, available using such devices as Kardia Band (AliveCor) and the Apple Watch.

The basic science principles behind these devices are the same as the traditional ECG. The device (whether it be a phone case, watch case, or other portable device) will have 2 metal plates that create the positive and negative electrodes of Lead I. When the right and left hands (or a wrist) touch both of these electrodes, a bipolar Lead I is created, as per Einthoven's original triangle (Figure 2) [17]. The signal is detected using the same principles of voltage conductance and vector analysis as the traditional ECG and interpreted using propriety AI software [18]. This ECG can then be stored, printed, or sent directly to physicians for interpretation and management.

**Figure 2.** (A) A photograph an Apple Watch series 4, an example of a wearable electrocardiogram device. The underside of the watch acts as the positive terminal, whilst the digital crown electrode acts as the negative terminal for Lead I (marked with + and −). When the user touches both simultaneously, a tracing from the view of Lead I can be recorded. (B) The second panel demonstrates the vector path this takes (RA to LA) on Einthoven's triangle. RA: Right arm; LA: Left arm; LL: Left leg




Ambulatory cardiac monitoring is by no means a new development; Holter first reported the use of his eponymous cardiac monitor in 1961 [19,20]. However, this new hardware represents a large step forward in making it more accessible and has several advantages over the traditional Holter monitor. Whilst portable, Holter monitors are still bulky and uncomfortable to wear; they require the patient to visit technicians for the placement and removal of electrodes; they are costly to health systems; they cannot be given to patients indefinitely; and they require patients to take the initial step of visiting a physician [19]. This is particularly important, as the asymptomatic patient unaware of their arrhythmia will not present until serious sequelae (eg, stroke secondary to AF) occur. Furthermore, patients are often monitored for 24-48 hours, which has been shown to miss up to 30% of clinically significant arrhythmias [21].

Undoubtedly, consumer-owned smart technology negates many of these limitations. The question of efficacy remains. One of the largest trials to date has been the Apple Heart Study, including detailed data for over 400 patients [5,18]. In this study, of the 400,000 initially recruited patients, over 2000 (0.5%) received a notification for irregular heart rate. Among patients with detailed data available, the positive predictive value was 0.84 (95% CI 0.76-0.92) for an irregular pulse notification detecting AF. Most studies are restricted to screening for AF, and a systematic review has observed overall sensitivities of around 94% and specificities of 93%-96%, depending on whether a smartphone or smartwatch was used [22].

Not only has the physical hardware become more portable and acceptable to patients, but the underlying software interpreting the acquired ECG has also improved drastically over recent years. Automated interpretations from traditional ECG machines have been reported as incorrect between 9% and 35% of interpretations; however, this depends on what rhythm is being evaluated (with AF being a particularly troublesome arrhythmia to diagnose) [23,24]. Newer smart-device AI can learn and adapt when exposed to a new "learning set" of patient results. By providing vast training sets of data to these algorithms in testing,

their overall efficacy is improved, compared with traditional ECG auto interpretation, which relies on applying strict measurement parameters to the ECG presented, without the capacity for learning [25]. For instance, in one of the seminal papers to describe this breakthrough, a learning set of 109 patients with AF was used, which resulted in the algorithm adjusting its weighting for P-wave absence [18,20]. This optimized algorithm had a sensitivity of 100% and a sensitivity of 96% compared with the initial values of 87% and 97%, respectively [18]. In an era of greater connectivity, the potential for crowdsourcing enormous datasets has resulted in more accurate and reliable algorithms, with several proprietary and open-source AF-detection algorithms available currently [25,26]. This demonstrates how deep learning that can now be used in real time for ECG analyses has the potential to far surpass previous automatic ECG interpretations.

## Wearable ECG Monitoring in Clinical Practice

The main use of these devices in clinical practice is the detection or exclusion of arrhythmias. KardiaPro has been approved in the United States for the screening and detection of AF, but has been studied in various other conditions including ventricular dysrhythmias, atrioventricular node re-entrant tachycardia, myocardial ischemia, and electrolyte disturbances [18,26-29]. AF is one of the most investigated applications as it is commonly asymptomatic, has a high prevalence (up to 1.4% of all patients aged >65 years), and can lead to devasting consequences such as stroke and death [30]. Studies examining the use of wearable ECG technology for screening of AF are broadly supportive; the SEARCH-AF Study used wearable ECG screening in pharmacies and found newly diagnosed AF in 15 patients (1.5%), with an overall prevalence of 6.7% [31]. A subsequent hypothetical community screening economic analysis extrapolated these results into a cost-effectiveness ratio of US $4066 per quality-adjusted life year gained, and a cost of US $20,695 for the prevention of 1 stroke [31]. When compared with the average inpatient costs of stroke (estimated at US $20,396 ± $23,256) plus associated outpatient costs (US $17,081 for the first-year plus US $16,689 for every year after), this represents potentially an enormous cost saving [32,33]. An Australian study using similar technology introduced nurse-led smartphone-based AF screening to general practices. The sensitivity and specificity of the automated algorithm were 95% (95% CI 83% - 99%) and 99% (95% CI 98% - 100%), respectively, and a new diagnosis of AF occurred in 0.8% of patients [34]. The evidence base for using these devices in screening at-risk populations is steadily increasing, and several further trials are planned for examining wearable ECG technology in other populations, including children [26,34,35]. Case reports exist of wearable ECG technology detecting cardiac ischemia [36] exercise-related arrhythmias in athletes [37], and polymorphic ventricular tachycardia [38], although these are not as commonly studied as the use of ECG for AF screening.

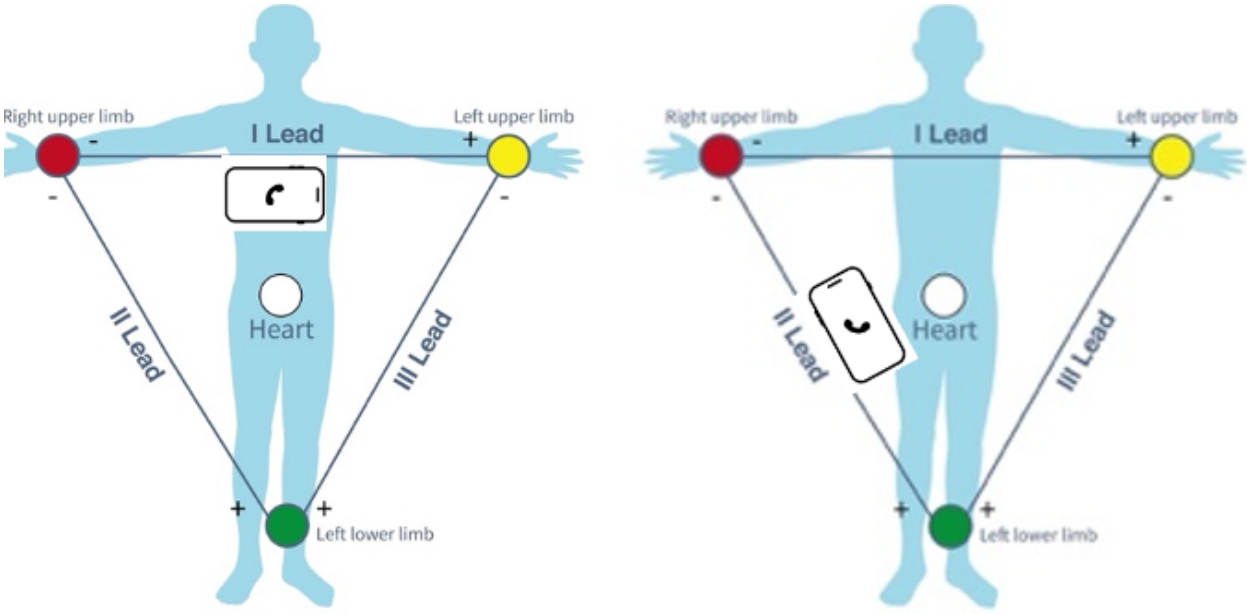
The reasons for these potential benefits over existing methodologies of AF screening and diagnosis have already been discussed; some of the biggest advantages are that patients are more likely to wear these comfortable, easily accessible devices, faster ECG analysis using AI algorithms with increasing diagnostic accuracy, and that data can be read in real time by physicians. There is also a health service economic incentive, as these devices can be bought by patients themselves for a fraction of the cost of a Holter monitor, at no cost to health systems and comparable efficacy for some dysrhythmias [5]. Patients themselves are also enthusiastic; a survey of 88 people showed that 82% found the device useful and the use of the device prompted a doctor's visit in 25% of patients [27]. While this obviously has a benefit if those patients did have arrhythmia, it does lead to questions surrounding resource use. This leads us to consider the potential limitations of this new technology.

## Limitations

This technology is not without its potential drawbacks to both the patient and the clinician. One of the largest technical drawbacks of this technology is its reliance using Lead I. Having only 1 positive and 1 negative electrode will only ever be able to provide a 1-lead view as the potential difference cannot be measured at further points (and thus obtain more leads) without more physical electrodes. It is not even possible to obtain augmented limb leads (which are unipolar and so could practically be created using only 1 positive electrode) as the neutral central terminal (Wilson's Central Terminal) is created by the average of Lead I, Lead II, and Lead III (3 leads). This can make the interpretation of dysrhythmias more difficult. For instance, having only 1 lead makes diagnosis of conduction delays like a right bundle branch block difficult as the characteristic pattern (rSR' in V1) is not necessarily visible in Lead I. Having only 1 lead on an extremity also increases the risk of artifacts; without other leads to compare with, artifactual "noise" is more difficult to exclude, and this noise can be amplified by having only 1 loosely attached electrode compared with traditionally several firmly attached electrodes.

One method of circumventing these limitations, however, is by changing the positioning of the positive terminal of the electrode (Figure 3). By keeping the negative terminal in the right hand and moving the positive terminal to the left leg, the potential difference being measured is in line with Lead II, providing now a 2-lead view of the heart. This has been shown to improve the diagnostic accuracy of some cardiac arrhythmias, especially atrial flutter, which may be more visible in inferior leads [39]. By simply moving this electrode, the sensitivity for atrial flutter increased from 27.3% to 72.7% [39].

**Figure 3.** Electrocardiogram vector change with repositioning. If the orientation of the phone is changed by repositioning the left-hand electrode to the right leg, the lead window changes from I to II.




There are other patient limitations. Using home ECG monitoring relies on patient technical skill set, as well as financial security to purchase one of these devices, and have consistent internet connectivity. With an aging population, the population that may benefit the most from the detection of occult arrhythmias (ie, older population) may be the group that struggles the most with adopting this technology. In addition, financial cost and consistent internet connectivity may also prove challenges for widespread adaptation.

The other major limitation is the practicality of physician access. Ironically, one of the greatest strengths of these devices (24-hour continuous monitoring for as long as the patient wants) can also be a weakness. Whilst a patient who has this technology now can record an ECG at any point in the day (or night), that does not necessarily mean that they will have timely access to a physician across the same hours. Patients who detect a possible arrhythmia outside of their doctor's availability may be left with 2 options: wait until an appointment becomes available, worrying all the while about potential strokes or cardiac events; or visit their nearest emergency department. From a resource use standpoint, this becomes worrisome, as in some studies, up to 7.3% of normal ECGs were reported as abnormal (sensitivity 97.1%, specificity 78.5%). Applied to the real world, that means 7 of every 100 normal ECGs may be reported as abnormal, resulting in 7 potentially unnecessary hospital visits per 100 normal ECGs. The question of what to do with patients who present with an abnormal ECG taken on a single lead private device is a vexing one. One potential solution could be rotating on-call physicians to review ECGs as they come through (as these can be sent in real time). However, this will leave open questions of compensation for the physician, and the eternal question raised above: how confident can a physician be based of a 1-lead ECG that there is no further pathology to exclude? What are the medicolegal implications of not fully working up a patient with a single positive trace who then has a devastating cardiovascular event? These issues need to be considered for the clinician to provide safe and sound medical treatment and advice to patients and as the prevalence of these devices rises, these are issues that will be faced by more and more clinicians.

Risk stratification may be useful here. The RITMO study examined whether having a higher screening threshold in elderly patients with hypertension and heart failure would increase AF capture rates. In this study, by stratifying by the stroke risk analysis algorithm, the rates of AF capture increased from the reported 3% at baseline to 13.2% [40]. By building risk stratification software into these devices, appropriate health care use could perhaps be improved.

Conversely, the lack of follow-up may be another limitation. Institution-provided monitors (eg, Holter monitors) have their data reviewed by physicians, and patient follow-up is initiated in the event of significant dysrhythmias. With consumer-owned devices, there is no assurance of follow-up, even if a significant arrhythmia is detected and the patient alerted. This has been borne out in real-life data, with only 57% of patients in the Apple Health Study with an irregular heart beat notification contacting healthcare providers [5].

## Conclusions

With an ever-growing health technology sector, wearable biometrics are more and more likely to appear outside of clinical research and into clinical practice. Although the machine taking the recordings becomes smaller and the software interpreting the readings becomes smarter, the underlying principles remain the same as what Einthoven first noticed some 100 years ago. If a clinician is then to have an informed discussion with a patient regarding the use of a wearable ECG device, then they must have confidence in their basic sciences to explain the mechanisms and potential limitations of such a device. With the anticipated explosion of these devices in people's private lives, questions surrounding this are almost a given, and thus,

all clinicians should be well acquainted with the basic sciences of electrocardiography.

Wearable ECG devices have many advantages over existing methods of trace acquisition, but also many potential drawbacks. The ease of use, patient-centered care, and increased availability of ECG monitoring must be balanced with a physician's duty of care and the potential for false-positive results, creating unnecessary unease and overtesting, as well as technical limitations of the devices themselves. Additional research and guidelines regarding the placement of a potential Lead II view, as well as thorough guidelines regarding data management, confidentiality, and physician workload need to be developed quickly before this technology becomes the standard.

## Acknowledgments

We thank Shutterstock illustrator asia11m for the use of the graphics (provided under license).

## Conflicts of Interest

None declared.

## References

1. Kligfield P. The centennial of the Einthoven electrocardiogram. J Electrocardiol 2002;35 Suppl:123-129. [doi: 10.1054/jelc.2002.37169] [Medline: 12539109]
2. Kligfield P, Gettes LS, Bailey JJ, et al. Recommendations for the standardization and interpretation of the electrocardiogram: part I: The electrocardiogram and its technology: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society: endorsed by the International Society for Computerized Electrocardiology. Circulation 2007 Mar 13;115(10):1306-1324. [doi: 10.1161/CIRCULATIONAHA.106.180200] [Medline: 17322457]
3. Rachim VP, Chung WY. Wearable noncontact armband for mobile ECG monitoring system. IEEE Trans Biomed Circuits Syst 2016 Dec;10(6):1112-1118. [doi: 10.1109/TBCAS.2016.2519523] [Medline: 27214910]
4. Lubitz SA, Faranesh AZ, Selvaggi C, et al. Detection of atrial fibrillation in a large population using wearable devices: the fitbit heart study. Circulation 2022 Nov 8;146(19):1415-1424. [doi: 10.1161/CIRCULATIONAHA.122.060291] [Medline: 36148649]
5. Perez MV, Mahaffey KW, Hedlin H, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019 Nov 14;381(20):1909-1917. [doi: 10.1056/NEJMoa1901183] [Medline: 31722151]
6. Zenzes M, Seba P, Portocarrero Vivero-Fäh B. The electrocardiogram on the wrist: a frightening experience to the untrained consumer: a case report. J Med Case Rep 2023 Mar 5;17(1):79. [doi: 10.1186/s13256-023-03806-3] [Medline: 36871070]
7. Galvani L. Aloysii galvani de viribus electricitatis in motu musculari commentarius. In: De Bononiensi Scientiarum et Artium Instituto Atque Academia 1791:363-418.
8. Raju TN. The Nobel Chronicles. 1924: Willem Einthoven (1860-1927). Lancet 1998 Aug;352(9139). [doi: 10.1016/S0140-6736(05)79240-0] [Medline: 9820341]
9. Barold SS. Willem Einthoven and the birth of clinical electrocardiography a hundred years ago. Card Electrophysiol Rev 2003 Jan;7(1):99-104. [doi: 10.1023/a:1023667812925] [Medline: 12766530]
10. Fisch C. Centennial of the string galvanometer and the electrocardiogram. J Am Coll Cardiol 2000 Nov 15;36(6):1737-1745. [doi: 10.1016/s0735-1097(00)00976-1] [Medline: 11092639]
11. Klabunde RE. Cardiac electrophysiology: normal and ischemic ionic currents and the ECG. Adv Physiol Educ 2017 Mar 1;41(1):29-37. [doi: 10.1152/advan.00105.2016] [Medline: 28143820]
12. Gulizia MM, Casolo G, Zuin G, et al. ANMCO/AIIC/SIT Consensus Information Document: definition, precision, and suitability of electrocardiographic signals of electrocardiographs, ergometry, Holter electrocardiogram, telemetry, and bedside monitoring systems. Eur Heart J Suppl 2017 May;19(Suppl D):D190-D211. [doi: 10.1093/eurheartj/sux031] [Medline: 28751842]
13. A R cardiac physiology: membrane potential, action potential, automaticity and electrical vectors. Clinical ECG interpretation. 2019. URL: https://ecgwaves.com/cardiac-electrophysiology-ecg-action-potential-automaticity-vector [accessed 2019-05-09]
14. Dupre A, Vincent S, Iaizzo PA. Basic ECG theory, recordings, and interpretation. In: Iaizzo PA, editor. Handbook of Cardiac Anatomy, Physiology, and Devices: Humana Press; 2005:191-201.
15. Becker DE. Fundamentals of electrocardiography interpretation. Anesth Prog 2006;53(2):53-63. [doi: 10.2344/0003-3006(2006)53[53:FOEI]2.0.CO;2] [Medline: 16863387]
16. Neubeck L, Lowres N, Benjamin EJ, Freedman SB, Coorey G, Redfern J. The mobile revolution--using smartphone apps to prevent cardiovascular disease. Nat Rev Cardiol 2015 Jun;12(6):350-360. [doi: 10.1038/nrcardio.2015.34] [Medline: 25801714]
17. Inc A. Taking an ECG with the ECG app on Apple Watch Series 4. 2019. URL: https://support.apple.com/en-us/HT208955 [accessed 2019-05-08]

18. Lau JK, Lowres N, Neubeck L, et al. iPhone ECG application for community screening to detect silent atrial fibrillation: a novel technology to prevent stroke. Int J Cardiol 2013 Apr 30;165(1):193-194. [doi: 10.1016/j.ijcard.2013.01.220] [Medline: 23465249]
19. Smith WM, Riddell F, Madon M, Gleva MJ. Comparison of diagnostic value using a small, single channel, P-wave centric sternal ECG monitoring patch with a standard 3-lead Holter system over 24 hours. Am Heart J 2017 Mar;185(1097-6744):67-73. [doi: 10.1016/j.ahj.2016.11.006] [Medline: 28267477]
20. HOLTER NJ. New method for heart studies. Science 1961 Oct 20;134(3486):1214-1220. [doi: 10.1126/science.134.3486.1214] [Medline: 13908591]
21. Lim YL, Mond H, Michael R, et al. Seven-day holter monitoring detects more significant arrhythmias than 24-hour and 3-day monitoring. Eur Heart J 2023 Jan 25;44(Supplement_1). [doi: 10.1093/eurheartj/ehac779.014]
22. Prasitlumkum N, Cheungpasitporn W, Chokesuwattanaskul A, et al. Diagnostic accuracy of smart gadgets/wearable devices in detecting atrial fibrillation: A systematic review and meta-analysis. Arch Cardiovasc Dis 2021 Jan;114(1):4-16. [doi: 10.1016/j.acvd.2020.05.015] [Medline: 32921618]
23. Anh D, Krishnan S, Bogun F. Accuracy of electrocardiogram interpretation by cardiologists in the setting of incorrect computer analysis. J Electrocardiol 2006 Jul;39(3):343-345. [doi: 10.1016/j.jelectrocard.2006.02.002] [Medline: 16777525]
24. Smulyan H. The computerized ECG: friend and foe. Am J Med 2019 Feb;132(2):153-160. [doi: 10.1016/j.amjmed.2018.08.025] [Medline: 30205084]
25. Bahrami Rad A, Kirsch M, Li Q, et al. A crowdsourced AI framework for atrial fibrillation detection in Apple watch and Kardia mobile ECGs. Sensors (Basel) 2024 Sep 2;24(17):5708. [doi: 10.3390/s24175708] [Medline: 39275619]
26. Barbagelata A, Bethea CF, Severance HW, et al. Smartphone ECG for evaluation of ST-segment elevation myocardial infarction (STEMI): Design of the ST LEUIS International Multicenter Study. J Electrocardiol 2018;51(2):260-264. [doi: 10.1016/j.jelectrocard.2017.10.011] [Medline: 29174099]
27. Orchard J, Lowres N, Freedman SB, et al. Screening for atrial fibrillation during influenza vaccinations by primary care nurses using A smartphone electrocardiograph (iECG): A feasibility study. Eur J Prev Cardiol 2016 Oct;23(2 suppl):13-20. [doi: 10.1177/2047487316670255] [Medline: 27892421]
28. Tabing A, Harrell TE, Romero S, Francisco G. Supraventricular tachycardia diagnosed by smartphone ECG. BMJ Case Rep 2017 Sep 11;2017:bcr2016217197. [doi: 10.1136/bcr-2016-217197] [Medline: 28899884]
29. Yasin OZ, Attia Z, Dillon JJ, et al. Noninvasive blood potassium measurement using signal-processed, single-lead ecg acquired from a handheld smartphone. J Electrocardiol 2017;50(5):620-625. [doi: 10.1016/j.jelectrocard.2017.06.008] [Medline: 28641860]
30. Freedman B, Camm J, Calkins H, et al. Screening for atrial fibrillation: a report of the AF-SCREEN international collaboration. Circulation 2017 May 9;135(19):1851-1867. [doi: 10.1161/CIRCULATIONAHA.116.026693] [Medline: 28483832]
31. Lowres N, Neubeck L, Salkeld G, et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies. The SEARCH-AF study. Thromb Haemost 2014 Jun;111(6):1167-1176. [doi: 10.1160/TH14-03-0231] [Medline: 24687081]
32. Godwin KM, Wasserman J, Ostwald SK. Cost associated with stroke: outpatient rehabilitative services and medication. Top Stroke Rehabil 2011 Oct;18 Suppl 1(sup1):676-684. [doi: 10.1310/tsr18s01-676] [Medline: 22120036]
33. Wang G, Zhang Z, Ayala C, Dunet DO, Fang J, George MG. Costs of hospitalization for stroke patients aged 18-64 years in the United States. J Stroke Cerebrovasc Dis 2014;23(5):861-868. [doi: 10.1016/j.jstrokecerebrovasdis.2013.07.017] [Medline: 23954598]
34. Gwynne K, Flaskas Y, O'Brien C, et al. Opportunistic screening to detect atrial fibrillation in Aboriginal adults in Australia. BMJ Open 2016 Nov 15;6(11):e013576. [doi: 10.1136/bmjopen-2016-013576] [Medline: 27852724]
35. Gropler MRF, Dalal AS, Van Hare GF, Silva JNA. Can smartphone wireless ECGs be used to accurately assess ECG intervals in pediatrics? A comparison of mobile health monitoring to standard 12-lead ECG. PLoS One 2018;13(9):e0204403. [doi: 10.1371/journal.pone.0204403] [Medline: 30260996]
36. Provencio A, Gil M. Smartwatch electrocardiogram records ST depression, premature ventricular complexes, and ventricular fibrillation. Lancet 2022 Nov 12;400(10364):e12. [doi: 10.1016/S0140-6736(22)01978-X] [Medline: 36366887]
37. Sanchis-Gomar F, Lavie CJ, Perez MV. Consumer wearable technologies to identify and monitor exercise-related arrhythmias in athletes. Curr Opin Cardiol 2021 Jan;36(1):10-16. [doi: 10.1097/HCO.0000000000000817] [Medline: 33074935]
38. Avidan Y, Tabachnikov V, Danon A, Schliamser JE. Polymorphic ventricular tachycardia detected by a smartwatch in a patient with recurrent syncope. JACC Case Rep 2024 Oct 2;29(19):102606. [doi: 10.1016/j.jaccas.2024.102606] [Medline: 39484328]
39. Rajakariar K, Koshy AN, Sajeev JK, Nair S, Roberts L, Teh AW. Modified positioning of a smartphone based single-lead electrocardiogram device improves detection of atrial flutter. J Electrocardiol 2018;51(5):884-888. [doi: 10.1016/j.jelectrocard.2018.07.008] [Medline: 30177334]
40. Andrade RD, Vitorino PVO, Sousa ALL, et al. A program to optimize the detection of paroxysmal atrial fibrillation: the RITMO study. Arq Bras Cardiol 2024;121(9):e20240235. [doi: 10.36660/abc.20240235] [Medline: 39352184]

## Abbreviations

**AF:** atrial fibrillation
**AI:** artificial intelligence
**ECG:** electrocardiogram

*Edited by A Coristine; submitted 29.05.24; peer-reviewed by I Jekova, I Iliev; revised version received 22.12.24; accepted 23.12.24; published 10.02.25.*

*Please cite as:*
*Smith S, Maisrikrod S*
*Wearable Electrocardiogram Technology: Help or Hindrance to the Modern Doctor?*
*JMIR Cardio 2025;9:e62719*
*URL: https://cardio.jmir.org/2025/1/e62719*
*doi:10.2196/62719*

# Applications of Ballistocardiogram in the Diagnosis of Coronary Heart Disease: Systematic Review

Mairihaba Maimaiti, MS; Hongling Zhu*, MD, PhD; Hesong Zeng*, MD, PhD

Department of Internal Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No. 1095 Jiefang Avenue, Wuhan, China

*these authors contributed equally

**Corresponding Author:**
Hesong Zeng, MD, PhD
Department of Internal Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No. 1095 Jiefang Avenue, Wuhan, China

## Abstract

**Background:** Coronary heart disease (CHD) continues to account for a substantial proportion of deaths worldwide. Ballistocardiogram (BCG), a noncontact, noninvasive technique for monitoring cardiac activity, has gained increasing attention for its potential role in various medical applications, particularly in CHD. This review comprehensively explores the applications of BCG in the diagnostic evaluation of CHD.

**Objective:** The aim of this systematic review is to evaluate the clinical applications and diagnostic capabilities of BCG in CHD, with the ultimate goal of enhancing the precision of CHD management and optimizing therapeutic decision-making pathways.

**Methods:** A literature search was performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines to identify studies evaluating the use of BCG in CHD. The initial search identified 500 studies. Based on titles, abstracts, and keywords, 266 studies were selected for further review. Following further exclusion of non-English articles, animal studies, and review articles, 38 eligible studies were included in the final analysis.

**Results:** Among the 38 studies, 22 focused on the application of BCG in acute coronary syndrome. These studies explored various aspects, including BCG waveforms in patients with acute myocardial infarction, the diagnosis of acute coronary syndrome, and the relationship between age and the rate of abnormal BCG waveforms. The remaining studies covered the effects of drugs, emotions, exercise, and other variables on BCG recordings in patients with CHD. Sample sizes varied significantly across the studies, 36 studies explicitly reported sample sizes, encompassing a total of 9479 participants with individual study sizes ranging from 1 to 903 cases. Notably, 13 studies enrolled fewer than 50 participants, raising concerns about potential selection bias and reduced reliability of the findings.

**Conclusions:** Overall, while BCG demonstrates significant potential in the diagnosis and prevention of CHD, several limitations remain. Variability in study design, sample size, and outcome measures poses challenges to the generalizability of findings. Nevertheless, the capability of BCG to reflect cardiac function and assist in the detection of CHD remains valuable. With continued research and technological advancement, BCG has the potential to transform current approaches to CHD diagnosis and management, ultimately improving patient outcomes and quality of life.

***(JMIR Cardio 2025;9:e68197)*** doi:10.2196/68197

**KEYWORDS**

ballistocardiogram; coronary heart disease; noninvasive; signal analysis; diagnostic performance; PRISMA

## Introduction

CHD remains a leading cause of global morbidity and mortality. It results from atherosclerosis and the subsequent narrowing or blockage of coronary arteries, leading to myocardial ischemia, hypoxia, or necrosis [1]. CHD may also involve other etiologies, such as inflammation and embolism, which can cause stenosis or occlusion of the vascular lumen. Despite the availability of diagnostic modalities, including electrocardiogram (ECG), exercise stress test, coronary computed tomography angiography, and coronary angiography, each method has inherent limitations. For instance, direct contact of ECG electrodes with the skin may cause allergic dermatitis. During acute myocardial infarction (MI) or severe arrhythmia, patients may be unable to perform the required exercise for the stress test. In addition, coronary computed tomography angiography involves a contrast agent injection, which may impair renal function or trigger allergic reactions, while also posing radiation risks. Although coronary angiography is the gold standard, it is invasive and costly. Consequently, developing a noninvasive, simple, and effective diagnostic method for CHD holds significant clinical value.

Research has definitively established that abnormal lipid metabolism is the primary pathogenic mechanism of CHD. Excess lipids accumulate on the arterial intima and penetrate the subendothelial space. Within the endothelium, these lipids are engulfed by macrophages, transforming into foam cells. As foam cells accumulate in significant numbers, they undergo apoptosis and necrosis, forming lipid necrosis cores. Over time, these cores evolve into atherosclerotic plaques, which gradually expand and narrow the vessels. When plaques rupture or ulcerate, thrombi may form, leading to blockage or severe stenosis of the coronary arteries [2].

BCG is a method that measures the force and velocity of the body's recoil, resulting from the ejection of blood from the heart during each heartbeat. This recoil force, known as the ballistic force, is detected by sensors placed on the body surface, typically on the trunk or limbs. BCG captures and analyzes the mechanical vibration signals generated by cardiac activity to infer coronary artery stenosis. In a healthy state, myocardial contraction and relaxation are coordinated and powerful, producing stable and regular vibration signals. However, when coronary artery stenosis occurs, myocardial blood supply is reduced, impairing contraction and relaxation [3,4]. This dysfunction alters the mechanical vibration signals of the heart, manifesting as waveform abnormalities, reduced amplitude, or changes in frequency.

Compared with traditional methods, the key advantage of BCG lies in its noninvasive and straightforward nature, avoiding allergic reactions and invasive procedures. Furthermore, it enables real-time, continuous monitoring of cardiac function, allowing clinicians to assess patient status and treatment response with greater precision. However, to validate BCG's value for CHD, further studies are needed to explore its correlation with biochemical indicators, such as blood lipid metabolism [5]. Such investigations will deepen our understanding of the potential and limitations of BCG in predicting coronary artery disease (CAD), thereby providing more accurate guidance for clinical diagnosis and patient management.

With the rapid advancement of artificial intelligence (AI) technologies, the processing and interpretation of BCG signals have also evolved toward more intelligent and automated approaches. Recent studies indicate that deep learning models, particularly convolutional neural networks and Siamese networks, offer promising solutions for feature extraction, anomaly detection, and disease assessment based on small-sample medical data [6,7]. These developments highlight the growing potential of integrating AI-integrated BCG analysis to enhance diagnostic accuracy and clinical utility.

This study systematically evaluates advancements and clinical applications of BCG specifically in CHD. We discuss BCG's potential in diagnosing and monitoring CHD, focusing on its role in patients with and without acute coronary syndrome (ACS). Although BCG shows great promise, its clinical use is still fragmented and lacks unified validation. This study aims to clarify BCG's diagnostic potential in CHD and explore how integrating AI technologies could improve diagnostic accuracy and expand its clinical applications.

The paper is structured as follows: Section 1 corresponds to the Introduction. Section 2 presents the Methods, focusing mainly on the search strategy. Section 3 contains the Results, which include the subsections: BCG in Detecting Medical Signals, BCG in ACS, and BCG in Non-ACS. Section 4 discusses the key findings and limitations of the study as well as the conclusions.

## Methods

### Literature Search

This review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines [8] and conducted a systematic search of English-language literature up to April 14, 2024, across PubMed, Scopus, Web of Science, and the Cochrane Library. The search strategy used keywords related to "ballistocardiogram," "ballistocardiograms," "ballistocardiograph," "ballistocardiography," "ballistocardiographic," "BCG," "coronary artery disease," "coronary atherosclerotic heart disease," "stable angina," "unstable angina," "ST segment elevation myocardial infarction," "non-ST segment elevation myocardial infarction." In addition, the reference lists of included studies were manually screened to identify further relevant literature. The full search strategies are provided in Multimedia Appendix 1.

### Selection Criteria

For the review, studies were included if they met the following criteria: (1) original research studies; (2) participants diagnosed with CHD; (3) research focus related to BCG; and (4) publication in English. Exclusion criteria were (1) non-English publications; (2) animal studies, reviews, or commentaries; (3) studies with unavailable full text; and (4) duplicate publications.

### Search Result

Studies were initially screened based on their titles, abstracts, and keywords, resulting in 266 potentially relevant studies. Of these, records were excluded due to inaccessible full texts (n=197), leaving records for full-text assessment (n=69). One reviewer performed the initial screening. For records with uncertain eligibility, a second reviewer was consulted to make the final inclusion decision. EndNote (Clarivate) was used for reference management, and no automation tools were applied in the screening process. Subsequently, non-English papers, reviews, and animal studies were excluded (n=17). Studies with duplicates were also removed (n=14). Finally, 38 studies [4,6,9-44] were included.

Data from the included studies were extracted by one reviewer. The extracted items included author and publication year, study content, sample size, and main findings. For any uncertainties during the extraction process, a second reviewer was consulted. No automation tools were used for data extraction, and the study authors were not contacted. EndNote was used solely for reference management. The extracted data are summarized in Multimedia Appendix 2.

Among the 38 studies [4,6,9-44], 22 [6,27,28,45-63] focused on the application of BCG in ACS. These studies explored various aspects, including BCG waveforms in patients with

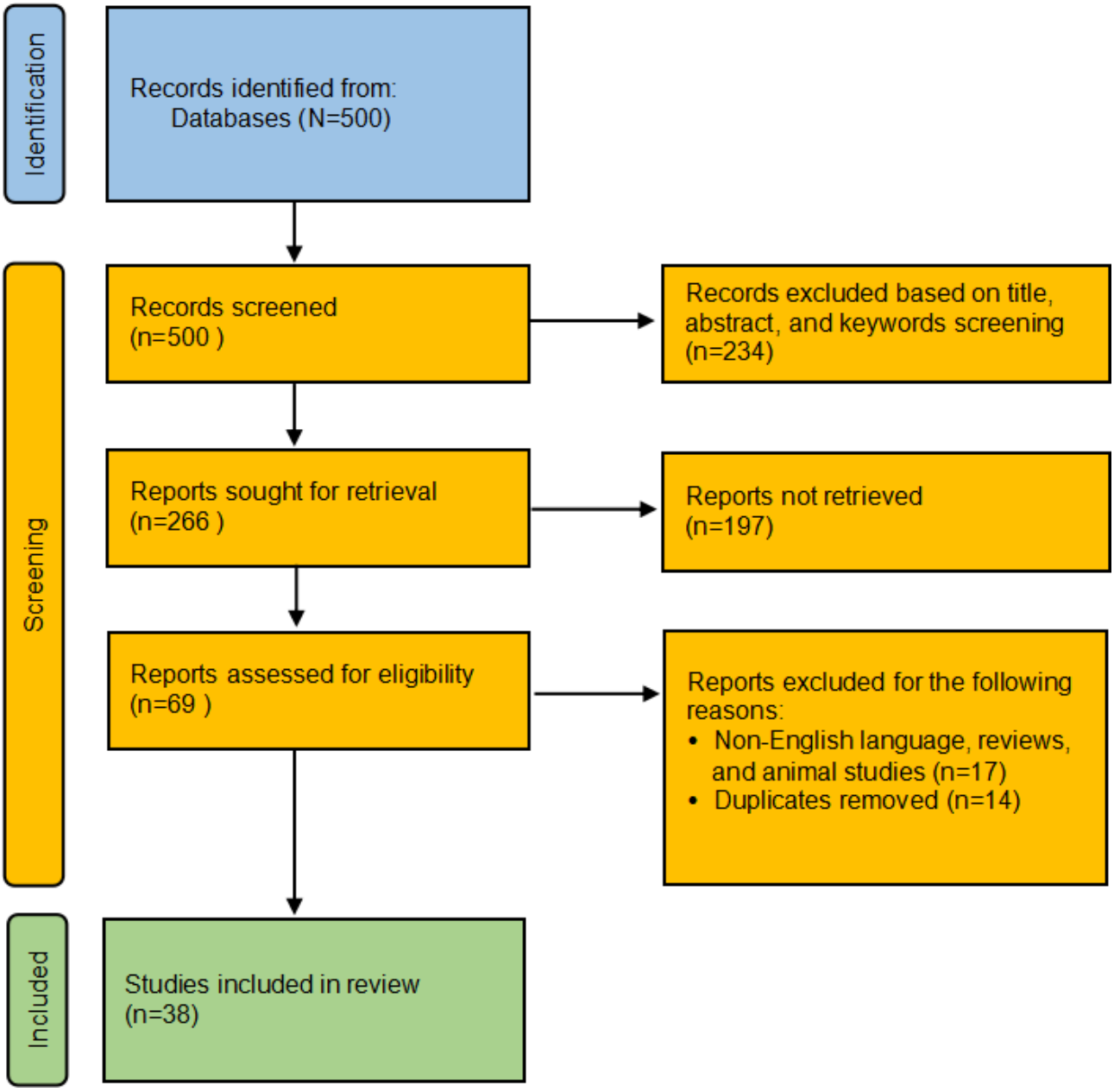
acute MI, the diagnosis of ACS, and the relationship between age and the incidence of abnormal BCG. The remaining studies covered the effects of drugs, emotions, exercise, and other variables on BCG recordings in patients with CHD.

The final selected studies were primarily conducted in the United States (n=27), the United Kingdom (n=4), China (n=3), Russia (n=1), and Sweden (n=3). Notably, the United States takes the lead, followed closely by the United Kingdom.

## Results

The flow diagram illustrated in Figure 1 presents the study selection process based on inclusion and exclusion criteria. Before presenting the results of the included studies, we briefly introduce the general applications of BCG in medical signal detection.

**Figure 1.** Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram.




### BCG in Detecting Medical Signals

Myocardial ischemia disrupts the heart's conduction system, causing abnormal electrical impulses and arrhythmias, such as tachycardia or bradycardia. Coronary atherosclerosis narrows the lumen, potentially reducing or obstructing blood flow to distal areas, leading to ischemia or necrosis. To compensate, the heart increases its beat frequency to enhance cardiac output and alleviate inadequate perfusion. However, this compensatory tachycardia raises myocardial oxygen demand, worsening ischemia and perpetuating a vicious cycle. Hence, heart rate (HR) control is crucial in the management of CHD.

In HR monitoring, BCG signals have attracted considerable interest due to their rich physiological information. Unsupervised clustering algorithms [64-68] have been used to identify intrinsic patterns, while continuous wavelet transforms [69-74] effectively capture their time-frequency characteristics.

Adaptive threshold algorithms [75-78] further enhance detection accuracy, and methods, such as the cepstral [45] and template matching [46,73], enhance HR feature extraction. The integration of machine learning has significantly advanced HR detection by enabling the automatic identification of relevant features from complex BCG signals, thereby expanding its clinical applications. Notably, reduced heart rate variability (HRV) due to autonomic dysfunction has been observed in cardiovascular conditions, such as myocardial ischemia, as well as in respiratory diseases like chronic obstructive pulmonary disease [47], highlighting the relevance of HRV analysis. In addition, the combination of beat-by-beat and BCG technology can effectively evaluate left ventricular function [9,10] and has the potential to become a noninvasive, low-cost left ventricular function assessment tool in the future. However, large-scale validation and rigorous methodological controls remain essential to establish its clinical reliability.

Autonomic neuropathy and vascular dysfunction frequently coexist and influence each other's manifestations. In cases where coronary artery stenosis results in myocardial ischemic necrosis, the damaged myocardium stimulates chemical and mechanical receptors on the cardiac wall, triggering a sympathetic stress response. This heightened sympathetic activity reduces HRV [48], thereby increasing the risk of cardiac instability. Therefore, effective management of CHD requires an integrated approach that addresses both autonomic and vascular components to optimize cardiac function.

The sophistication of BCG signals for HRV detection has enabled diverse applications. Sensors like ElectroMechanical Film [49-51] and Polyvinylidene Fluoride [53-55], commonly embedded in seat belts and smart beds [51,56,79], are widely used to capture BCG signals for HRV assessment. In addition, physiological signals can be recorded using data acquisition systems such as BIOPAC (BIOPAC Systems) [52]. These technologies are adaptable to various environments, expanding the scope of HRV monitoring. Comparative studies highlight BCG's utility, showing strong agreement between BCG and ECG in HRV measurements and revealing consistent trends [50]. This concordance positions BCG as a valuable complement to ECG, enhancing the comprehensiveness of HRV assessment. With ongoing technological advances and further research, the applications of BCG in HRV detection are expected to expand. Its versatility and reliability in various settings promise a bright future for BCG in HRV detection. Meanwhile, studies on wearable blood pressure (BP) monitoring, such as those using cuffless pulse transit time (PTT) methods, have demonstrated promising accuracy and usability, indicating the growing potential of wearable technologies in cardiovascular health monitoring [57].

Patients with CHD often exhibit complex and variable BP patterns. Before the onset of cardiac symptoms, BP may elevate, reflecting reduced cardiac compliance—diminished heart elasticity in response to blood flow, necessitating higher pressures for normal circulation [58]. However, during acute attacks, ischemia can significantly reduce myocardial contractile force, causing BP to drop. Compounding this, many patients also have hypertension, a chronic condition that not only accelerates coronary atherosclerosis but also chronically overloads the heart, further compromising myocardial function. This dual disease burden significantly raises the risk of cardiovascular events, namely MI and heart failure.

BCG has emerged as a standout in BP detection, leveraging its unique characteristics. The signal, generated by cardiac pulsations and arterial blood flow, represents changes in the external body pressure, enabling noncontact monitoring of heart activity. BCG, when combined with methods such as PTT or photoplethysmography, shows comparable accuracy to traditional cuff-based blood pressure devices. For instance, one study reported root mean square errors of 6.7 (SD 1.6) mmHg for systolic and 4.8 (SD 1.5) mmHg for diastolic BP, similar to values from standard sphygmomanometry [59,60]. By using techniques, such as clustering [61], regression analysis [62], and deep neural networks [80], key metrics like pulse arrival time [50,62,81] and PTT [63,82-91] can be accurately calculated to estimate BP. In addition, technological advancements are driving innovations in BCG-based BP monitoring methods. These include novel devices, such as chair-based systems [92,93] and wearable limb BCG [87], broadening the possibilities for accurate and convenient BP monitoring.

In summary, BCG offers distinctive advantages in detecting vital indicators. Its noninvasive and user-friendly nature positions it as a pivotal tool for cardiac function assessment, affording doctors greater precision and depth in diagnostic analysis.

## BCG in ACS

For patients with MI, especially those with non-ST-segment elevation MI and ST-segment elevation MI, the application of BCG technology holds significant importance. During an MI, myocardial cell necrosis and injury affect the heart's mechanical and electrical activities, leading to subtle waveform changes in BCG recordings. For instance, features such as a low I wave, deep K wave, high H wave, and a notched J wave can all serve as important indicators of MI [11-16]. Moreover, abnormal waveforms and amplitude changes are more pronounced on the acceleration curve than on the displacement or velocity curve [17]. A study [18] conducted thorough ECG and BCG measurements on 78 patients with angina pectoris and 32 patients with distal MI. The analysis revealed that 40 (36%) cases of patients exhibited abnormal ECGs, a concerning figure indeed. However, the most startling finding was that 88 (80%) cases showed abnormal BCG results, indicating high sensitivity of this diagnostic tool in detecting heart diseases. Furthermore, they discovered that 94 (85%) patients exhibited at least 1 or 2 abnormalities in either their ECGs or BCG, emphasizing the intricate and diverse nature of heart disease manifestations. Research [19-21] has shown that abnormal BCG can be recorded in MI. Concurrently, in one observational study [94], all patients with symptomatic MI exhibited BCG abnormalities, further demonstrating its potential in MI detection [30].

With increasing age, the incidence of abnormal ECG and BCG findings among older patients with MI shows a pronounced upward trend [22]. Specifically, the frequency of such abnormalities increases significantly with each decade, particularly when compared to age-matched healthy individuals. This significant finding underscores the unique cardiac function

changes in older patients with MI, holding immense importance for the prevention and early detection of this condition. Studies have reported that clinically healthy individuals under 40 years of age rarely exhibit abnormal BCG recordings. However, after the age of 80 years, the occurrence of abnormal electrocardiographic findings rises sharply, approaching 100% in certain cohorts [23]. This escalating trend underscores the deterioration of central organ function and the escalating risk of heart disease among the older population, emphasizing the need for vigilant monitoring and proactive management. Cardiovascular aging is a natural phenomenon that involves the gradual deterioration of vascular structure and function as a person ages [95]. A study used BCG to observe 21 patients under 54 years old with a history of MI. They found that 17 (81%) cases exhibited abnormal BCG results, directly reflecting accelerated cardiovascular aging, particularly a significant decrease in HJ force (the force of the H and J waves of the BCG) [24]. This abnormality rate was significantly higher than that in age-matched healthy controls, indicating signs of accelerated cardiovascular aging. These findings suggest that BCG-derived metrics may serve as early indicators of subclinical ischemic heart disease (IHD), allowing for earlier risk stratification and preventive intervention in at-risk populations.

Portable BCG monitoring has demonstrated potential as a diagnostic tool for detecting CAD in patients with angina pectoris [25]. In cases of insufficient coronary artery blood supply, there is a high incidence of abnormal BCG patterns, with the deep K stroke pattern occurring frequently. A parallel study [5] investigated the correlation between BCG and MI, as well as the interplay with blood lipids. The findings suggested that BCG may serve as an indicator of abnormalities in individuals predisposed to MI and that BCG abnormalities may be influenced by factors, such as smoking, physical exertion, and emotional state. Meanwhile, BCG analysis incorporating respiratory monitoring successfully identified left ventricular dysfunction beats in a United States cohort study [10]. Using a nonlinear quadratic discriminant function, the study discovered that 87% (239/275) of the heartbeats in healthy males were classified as “normal,” while 98% (45/46) of the heartbeats in males with CHD were accurately identified as “coronary heart” beats. Building on prior research, a beat-by-beat analysis was conducted to assess left ventricular contractility and abnormal beats, ultimately discovering that 96% of heartbeats stemming from 6 patients with MI were accurately categorized as resembling “CAD-like” heartbeats [9], indicating the promising role of BCG in cardiac abnormality detection.

To better quantify and assess the severity of abnormalities in BCG after MI, some studies have introduced a grading system [16]. This system divides BCG into 4 grades: grade I represents minimal abnormalities, indicating that the patient’s heart function is basically normal; grade II indicates moderate abnormalities, suggesting that the patient’s heart function has recovered to some extent but is not yet fully normalized; grade III represents significant abnormalities, indicating severe damage to the heart with limited recovery; and grade IV represents the most severe abnormalities, suggesting severe impairment and poor recovery of heart function. This grading system provides clinicians a systematic framework for severity stratification, thereby informing more personalized and targeted treatment strategies.

BCG is not only used for the diagnosis of MI but also effectively assesses the treatment outcomes and prognosis of patients [26]. Research has shown that the prediction of subsequent MI or sudden death by BCG is highly accurate, with a statistical significance of $P<.001$ [27]. By comparing changes in BCG signals before and after treatment, clinicians can visually assess the recovery of patients’ heart function, predict potential risks of complications, and accordingly adjust and optimize treatment plans. With its real-time monitoring, high sensitivity, and specificity, BCG provides powerful support for the diagnosis and treatment of patients with MI.

## BCG in Non-ACS

A study using low-frequency BCG to assess cardiac function in patients with atherosclerotic heart disease found that, although IJ amplitude (the amplitude difference between the I wave and the J wave in the BCG signal) exhibited respiratory-related fluctuations in these patients, the changes were not significantly different from those observed in age-matched healthy individuals [28]. This finding contributes to our understanding of the physiological implications of arterial sclerosis, particularly regarding IJ amplitude variability. However, a pivotal observation emerged from their study. In cases where therapeutic interventions fail to arrest disease progression, a discernible upward trend in IJ amplitude becomes evident. This telling shift may serve as a harbinger of waning heart function or exacerbating pathological insults, thereby providing clinicians with a vital tool for evaluating treatment responsiveness and tracking disease trajectories.

Recent studies have expanded BCG’s diagnostic applications in cardiovascular assessment. Analysis of myocardial functional integrity [29] demonstrated particularly high clinical utility for life insurance evaluations. The research underscored the heightened prevalence of BCG abnormalities among the older populations. Subsequent studies [3] further validated BCG’s diagnostic prowess by using it to screen for suspected CHD in individuals with chronic chest pain, contrasting results with ECG findings. Among 197 patients with CHD, 159 (81%) exhibited an abnormal BCG, while only 14 (7%) cases presented with concurrent normal BCG and ECG readings, underscoring BCG’s diagnostic sensitivity. In another investigation, comparative analysis of BCG recordings from patients with IHD and healthy participants revealed robust correlations between IJ velocity, vanillylmandelic acid excess, and cardiovascular risk. Specifically, patients with IHD on the brink of mortality consistently exhibit lower initial IJ amplitudes, offering invaluable insights into the prognostic value of BCG in assessing imminent cardiac events.

Noninvasive BCG assessment achieved a diagnostic accuracy of 77%, correctly stratifying 289 of 375 patients with CHD severity when validated against the gold standard of coronary angiography [31]. This result highlights BCG’s promise as a diagnostic adjunct in cardiovascular assessment. A study [32] used BCG technology to monitor HR and respiration in the 3 months following coronary artery bypass grafting surgery. The findings revealed that the mean respiration rate was 21.8 (SD

2.5) breaths per minute, while the mean HR was 67.6 (SD 2.4) beats per minute. These findings underscore the stability and recovery trajectory of patients undergoing this complex surgical procedure. Building upon these successes, other researchers used BCG to assess myocardial contractility and prognosis before and after coronary artery bypass grafting [33]. Their findings indicated a noteworthy 3% increase in average HR and myocardial strength following the surgical intervention. This improvement not only validates the efficacy of the surgical intervention but also highlights the potential of BCG as a sensitive tool for assessing cardiac function and predicting patient outcomes.

An observational study compared the BCG of 77 patients with CHD and 48 healthy individuals and further analyzed the changes in BCG parameters, such as time interval from I wave to J wave, time interval from J wave to K wave, and energy of the HIJK wave complex, before and after surgery [34]. Notably, statistical analysis revealed significant differences across these indicators: time interval from I wave to J wave (mean 95, SD 9 vs mean 78, SD 11 ms), time interval from J wave to K (mean 72, SD 10 vs mean 63, SD 8 ms), and energy of the HIJK wave complex (mean 0.020, SD 0.009 vs mean 0.010, SD 0.006 $V^2$). Postoperatively, the BCG amplitude exhibited an increase, and the I-peak became significantly deeper, suggesting positive physiological changes. Also, a novel method was developed using micromotion-sensitive mattresses to gather BCG signals and then leveraged the ensemble empirical mode decomposition method to calculate HRV for disease classification [35]. This approach achieved a diagnostic accuracy of 92% by correctly classifying 17 of 18 participants. These results highlight the value of integrating advanced signal acquisition and processing in BCG-based diagnostics. Further advancing the field, another study [36] used the combination of short-time Fourier transform and ensemble empirical mode decomposition to accurately identify IJK complexes within BCG signals and assess HR. The proposed approach achieved a mean absolute error of 0.99 (95% CI −1.81 to 3.79) bpm, underscoring the high degree of accuracy and reliability achieved through the fusion of these advanced analytical techniques, opening new avenues for BCG-based cardiac monitoring and diagnosis.

The research has revealed that inducing hypoxemia serves as an advantageous approach in assessing critical physiological indicators like cardiac output and pulse pressure in humans under stressful conditions. This technique enables the maintenance of a consistent stress level for an extended period, ultimately ensuring the reliability and precision of collected data. Critically, the induction of hypoxemia minimizes the likelihood of artificial distortions in BCG recordings, a pivotal factor in securing accurate diagnostic outcomes [96]. As a result, when ECG and BCG readings in the resting state are inconclusive or ambiguous [37], it is advisable to consider capturing these 2 data types again in a simulated hypoxemic environment. This strategy may yield more profound and definitive diagnostic insights into CAD [38], particularly in scenarios where traditional diagnostic methods fail to provide clear-cut results. The value of this approach cannot be overstated.

As is well known, smoking history is one of the risk factors for CHD. Smoking damages endothelium and promotes atherosclerosis. Notably, smoking can also significantly alter the BCG of patients [39]. Studies have highlighted a stark contrast, with only approximately 8/114 (6.8%) healthy individuals experiencing BCG deterioration post smoking, compared to a staggering 51/86 (59%) individuals with CHD [40]. To further explore the intricate interplay between emotional states and BCG dynamics, BCG monitoring was performed on 48 patients with IHD [41], aiming to elucidate the association between emotional states and alterations in IJ velocity. The analysis demonstrated that 5/6 participants exhibited a positive correlation between IJ velocity and emotional arousal ($r$=0.34 - 0.90), while 4/6 participants displayed a positive correlation between HR and emotional arousal ($r$=0.60 - 0.90). These findings underscore the intricate link between emotional states and BCG parameters, particularly in patients with IHD, highlighting the potential clinical significance of monitoring BCG in such contexts.

Excessive exercise load can significantly exacerbate symptoms in patients with CHD and even trigger severe angina. In this emergency situation, nitroglycerin becomes the key medication to alleviate symptoms. According to research results, the majority of patients with CHD showed significant improvement in their BCG after taking nitroglycerin [13,42]. Nitrite also has an impact on the BCG. After inhaling amyl nitrite, healthy individuals may experience a temporary increase in BCG amplitude. In patients with CAD, this change is minimal. Moreover, the IJ waveform of BCG cannot accurately reflect the cardiac output [43]. BCG monitoring of isosorbide dinitrate therapy in patients with CHD demonstrated that normalization of BCG was more prevalent among patients experiencing symptom improvement compared to those with normal ECG [44]. This suggests that BCG provides a more sensitive and objective evaluation tool than ECG for assessing therapeutic effects in patients with CHD.

By continuously monitoring changes in heart electrical activity, BCG technology provides timely diagnostic evidence for doctors, enabling them to quickly adopt emergency treatment measures, effectively preventing further deterioration of the condition, and buying precious treatment time for patients.

## Discussion

### Principal Findings

This study provides a comprehensive review and analysis of the widespread application of BCG in the field of CHD and its use in monitoring vital signs, emphasizing its importance and future potential. BCG, as a novel technology for cardiovascular function monitoring, uses relevant techniques to detect key indicators such as HR, HRV, and BP. It not only offers a new method for HR monitoring but also represents a low-cost, high-precision, and noncontact technology. In CHD, BCG aids in diagnosis and disease classification, and it plays a critical role in assessing treatment effectiveness. According to the literature, 208/239 (87%) heartbeats from healthy males were correctly classified as “normal” through BCG analysis, while 45/46 (98%) heartbeats from male patients with CHD were

accurately identified as “CHD” heartbeats. When compared to coronary angiography, BCG showed a 77% accuracy by categorizing 289/375 patients according to their CHD severity. The application of BCG technology not only improves the accuracy and efficiency of diagnosis but also significantly enhances targeted and personalized treatment in patient care, revolutionizing the management and treatment of CHD.

### Limitations

In terms of the review process, several limitations should be noted. First, although we adopted a systematic approach, only English-language literature was included, which may introduce language bias. Second, data extraction was performed by a single reviewer, with a second reviewer consulted only when uncertainties arose; this could lead to potential subjectivity. Furthermore, no formal quality assessment of the included studies was conducted, and the review protocol was not prospectively registered, which may affect the transparency and reproducibility of the review.

Although BCG has become a research hotspot due to its potential and prospects in the medical field, its application in CHD research is still relatively limited. Most of the studies are concentrated in high-income countries, such as the United States and the United Kingdom. However, the total number of related studies remains small globally, indicating that the research activity and depth in this field remain insufficient. Although some studies have reported sample sizes exceeding 900, this number still appears relatively small compared to other more mature research fields, indicating that large-scale studies based on BCG are still significantly insufficient in CHD research. In terms of research content, the current focus is mainly on the analysis of waveform abnormalities, and there is a lack of in-depth research and clear conclusions on how to specifically use BCG to diagnose CHD, as well as which specific waveform changes correspond to CHD. This has limited the application of BCG technology in the diagnosis of CHD. As for accuracy, although BCG technology has shown an acceptable level of accuracy in current research, it is important to acknowledge that sensor design and signal processing itself is a complex and highly specialized field. These factors may all have an impact on the accuracy and reliability of BCG. Therefore, more in-depth and systematic research and validation are needed before applying it to the diagnosis of CHD. To improve BCG accuracy in CHD research, identifying and removing artifacts is crucial. Recent studies have highlighted the importance of multimodal monitoring, particularly the integration of ECG and BCG, for artifact reduction, offering valuable insights for BCG’s CHD application [97]. In addition, in home environments, BCG signals may be affected by data loss and motion artifacts due to patient movement or device misplacement, further affecting signal quality and diagnostic reliability [47]. In summary, although BCG has shown certain potential and prospects, there are still many challenges and limitations in its application in the diagnosis of CHD. Therefore, further research and development are required to better explore the potential of BCG and to establish a more robust and reliable theoretical and practical foundation for its application in the diagnosis of CHD. While this review focuses on the current clinical applications and challenges of BCG in CHD, the rapid advancement of AI presents an exciting opportunity to enhance BCG signal analysis. Future studies should investigate AI-driven feature extraction and predictive modeling to improve diagnostic accuracy and facilitate personalized cardiovascular monitoring.

### Conclusions

BCG is a noninvasive cardiovascular detection method. Through summarizing relevant literature, the application of BCG in CHD has been explored, mainly in recording BCG in patients with MI and angina pectoris. These studies have laid the foundation for the diagnosis and treatment of CHD in the future. With the continuous innovation and improvement of technology, BCG is expected to play an increasingly important role in the diagnosis and treatment of CHD. Its unique advantages and potential make it an important pillar in this field, bringing more hope and good news to patients. It is anticipated that as research deepens and technology matures, BCG will be able to diagnose more accurately, providing doctors with more comprehensive and accurate patient information, which will facilitate the development of personalized treatment plans. This will help improve the treatment effectiveness of CHD, enhance patients’ quality of life, and even potentially reduce medical costs, bringing greater benefits to society. At the same time, it is hoped that more researchers will contribute to the development of BCG, advancing progress and innovation in this field. Through interdisciplinary collaboration, further exploration of the potential of BCG in CHD diagnosis and treatment can be achieved. This will not only help promote the advancement of medical technology but also bring more blessings and hope to patients. In short, as an emerging technology, BCG has unlimited possibilities and hopes for application in the field of CHD. It is expected to improve treatment outcomes and quality of life for patients, while playing an increasingly important role in the development of the medical field.

### Acknowledgments

H Zeng and H Zhu are recognized as co-corresponding authors of this review for their leadership, critical input, and overall supervision during the preparation and revision of the manuscript. This work is funded by the National Natural Science Foundation of China (82100531), National Key Research and Development Program (2022YFC2407000, 2022YFC2407003), Tongji Hospital Medical Innovation and Translational Incubation Project (2023CXZH004), Medical Artificial Intelligence Fund of Tongji Hospital (AI2024A02), Tongji Hospital Excellent Young Scientist Fund (24-2KYC13057-14), and Huazhong University of Science and Technology's 2024 "Interdisciplinary Research Support Program" Project (2024JCYJ066-PP).

## Authors' Contributions

MM performed literature retrieval, data collection, and article screening. H Zhu verified ambiguous studies and made final inclusion decisions. MM and H Zhu prepared the initial draft. H Zhu and H Zeng critically revised the manuscript and supervised the research.

H Zhu is the co-corresponding author and can be reached via email at hlzhu0826@126.com and by phone at 86 159 2717 0295.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete search strategies for all databases.
[DOC File, 26 KB - cardio_v9i1e68197_app1.doc ]

Multimedia Appendix 2
Summary of included studies in the review.
[DOC File, 246 KB - cardio_v9i1e68197_app2.doc ]

Checklist 1
PRISMA checklist
[DOCX File, 274 KB - cardio_v9i1e68197_app3.docx ]

## References

1. Agrawal S, Mastana S. Genetics of coronary heart disease with reference to ApoAI-CIII-AIV gene region. World J Cardiol 2014 Aug 26;6(8):755-763. [doi: 10.4330/wjc.v6.i8.755] [Medline: 25228954]
2. Henderson A. Coronary heart disease: overview. Lancet 1996 Nov;348:S1-S4. [doi: 10.1016/s0140-6736(96)98001-0] [Medline: 8918518]
3. Wade EG, Fulton RM, Mackinnon J. The ballistocardiogram in the diagnosis of coronary arterial disease. Br Heart J 1956 Jan;18(1):65-77. [doi: 10.1136/hrt.18.1.65] [Medline: 13284187]
4. Baker BM. Ballistocardiography: predictor of coronary heart disease. Circulation 1968 Jan;37(1):1-3. [doi: 10.1161/01.cir.37.1.1] [Medline: 5634725]
5. Smith DH. The ballistocardiogram as a predictor of myocardial infarction. Am J Cardiol 1959 Feb;3(2):247-249. [doi: 10.1016/0002-9149(59)90295-4] [Medline: 13626856]
6. Abdulraheem M, Oladipo ID, Ajagbe SA, Balogun GB, Emma-Adamah NO. Continuous eye disease severity evaluation system using Siamese neural networks. paradigmplus 2023 Mar;4(1):1-17. [doi: 10.55969/paradigmplus.v4n1a1]
7. Ajagbe SA, Adigun MO. Deep learning techniques for detection and prediction of pandemic diseases: a systematic literature review. Multimed Tools Appl 2023 May 29;29(1-35):1-35. [doi: 10.1007/s11042-023-15805-z] [Medline: 37362693]
8. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
9. Dinaburg AG, Zuckerman BR. Components of left ventricular dysfunction discriminated noninvasively beat by beat. Am J Cardiol 1987 Nov 1;60(13):1123-1127. [doi: 10.1016/0002-9149(87)90365-1] [Medline: 3673911]
10. Dinaburg AG, Zuckerman BR. Left ventricular dysfunction discriminated noninvasively beat by beat. Am J Cardiol 1984 Jan 1;53(1):238-242. [doi: 10.1016/0002-9149(84)90719-7] [Medline: 6691267]
11. Moser M, Pordy L, Chesky K, Taymor RC, Master AM. The ballistocardiogram in myocardial infarction: a study of one hundred cases. Circulation 1952 Sep;6(3):402-407. [doi: 10.1161/01.cir.6.3.402] [Medline: 14954536]
12. Kuo PT, Joyner CR Jr. Angina pectoris induced by fat ingestion in patients with coronary artery disease; ballistocardiographic and electrocardiographic findings. J Am Med Assoc 1955 Jul 23;158(12):1008-1013. [doi: 10.1001/jama.1955.02960120008004] [Medline: 14392044]
13. Starr I, Pedersen E, Corbascio AN. The effect of nitroglycerine on the ballistocardiogram of persons with and without clinical evidence of coronary heart disease. Circulation 1955 Oct;12(4):588-603. [doi: 10.1161/01.cir.12.4.588] [Medline: 13261311]
14. Isaacs J, Wilburne M, Gunther L. Ultra-low frequency acceleration (force) ballistocardiogram in myocardial infarction. Am J Cardiol 1958 Mar;1(3):323-333. [doi: 10.1016/0002-9149(58)90297-2] [Medline: 13508554]
15. de Faire U, Theorell T. A simple ballistocardiographic measure (IJ/HI amplitude ratio) in relation to electrocardiographic evidence of ischaemic heart disease. Scand J Clin Lab Invest 1979 Sep;39(5):435-440. [doi: 10.3109/00365517909106128] [Medline: 574981]

16. Mandelbaum H, Mandelbaum RA. Studies utilizing the portable electromagnetic ballistocardiograph. IV. The clinical significance of serial ballistocardiograms following acute myocardial infarction. Circulation 1953 Jun;7(6):910-915. [doi: 10.1161/01.cir.7.6.910] [Medline: 13051833]
17. Smith JE. Comparison of the displacement, velocity, and acceleration ballistocardiography in coronary heart disease. Am Heart J 1953 Nov;46(5):692-704. [doi: 10.1016/0002-8703(53)90222-1] [Medline: 13104313]
18. Baker BM, Scarborough WR, Mason RE, Singewald ML. Coronary artery disease studied by ballistocardiography: a comparison of abnormal ballistocardiograms and electrocardiograms. Trans Am Clin Climatol Assoc 1950;62:191-201. [Medline: 21407742]
19. Rivin AU. Clinical experience with the portable electromagnetic ballistocardiograph. Calif Med 1954 Jan;80(1):16-20. [Medline: 13116023]
20. Harrison WK, Talbot SA. Discrimination of the quantitative ultralow-frequency ballistocardiogram in coronary heart disease. Am Heart J 1967 Jul;74(1):80-87. [doi: 10.1016/0002-8703(67)90044-0] [Medline: 6027574]
21. Taymor RC, Pordy L, Chesky K, Moser M, Master AM. The ballistocardiogram in coronary artery disease. JAMA 1952 Feb 9;148(6):419. [doi: 10.1001/jama.1952.02930060001001]
22. Scarborough WR, Mason RE, Davis FW Jr, Singewald ML, Baker BM Jr, Lore SA. A ballistocardiographic and electrocardiographic study of 328 patients with coronary artery disease; comparison with results from similar study of apparently normal persons. Am Heart J 1952 Nov;44(5):645-670. [doi: 10.1016/0002-8703(52)90095-1] [Medline: 12985532]
23. Scarborough WR, Smith EW, Baker BM. Studies on subjects with and without coronary heart disease. Serum lipid, lipoprotein, and protein determinations and their relation to ballistocardiographic findings (a preliminary survey). Am Heart J 1960 Jan;59:19-35. [doi: 10.1016/0002-8703(60)90381-1] [Medline: 14442227]
24. Moss AJ. Ischemic heart disease and accelerated cardiovascular aging. A ballistocardiographic study. Circulation 1962 Feb;25:369-375. [doi: 10.1161/01.cir.25.2.369] [Medline: 14476526]
25. Mandelbaum H, Mandelbaum RA. Studies utilizing the portable electromagnetic ballistocardiograph. I. Abnormal HIJK patterns in hypertensive and coronary artery heart disease. Circulation 1951 May;3(5):663-673. [doi: 10.1161/01.cir.3.5.663] [Medline: 14831188]
26. Theorell T, Rahe CR. Life change events, ballistocardiography and coronary death. J Human Stress 1975 Sep;1(3):18-24. [doi: 10.1080/0097840X.1975.9939543] [Medline: 1235115]
27. Lynn TN, Wolf S. The prognostic significance of the ballistocardiogram in ischemic heart disease. Am Heart J 1974 Sep;88(3):277-280. [doi: 10.1016/0002-8703(74)90459-1] [Medline: 4855021]
28. Jones RJ. The Nickerson ballistocardiogram in arteriosclerotic heart disease with and without congestive failure. Circulation 1952 Sep;6(3):389-401. [doi: 10.1161/01.cir.6.3.389] [Medline: 14954535]
29. Ungerleider HE, Gubner R. Application of heart function tests in life insurance medicine. J Intern Med 1953 Jan 12;145(s277):45-51. [doi: 10.1111/j.0954-6820.1953.tb04065.x]
30. Theorell T, Blunk D, Wolf S. Ballistocardiographic indicators of prognosis in ischemic heart disease. J Lab Clin Med 1975 Jul;86(1):46-56. [Medline: 1151142]
31. Bancroft WH, Swatzell RH, Baldone JC, Tucker MS, Eddleman EE. Computerized system for evaluation of coronary artery disease by noninvasive techniques. Presented at: ACMSE '76: the 14th Annual ACM Southeast Regional Conference; Apr 22-24, 1976; Birmingham, Alabama. [doi: 10.1145/503636.503640]
32. Harrington N, Bui QM, Wei Z, et al. Passive longitudinal weight and cardiopulmonary monitoring in the home bed. Sci Rep 2021 Dec 21;11(1):24376. [doi: 10.1038/s41598-021-03105-1] [Medline: 34934065]
33. Starr I, MacVaugh H III. Early and late effects of the coronary bypass operation on cardiac contractility and coordination. Am Heart J 1975 Aug;90(2):179-189. [doi: 10.1016/0002-8703(75)90118-0]
34. Cao X, Liu L, Cai D, Guo P, Tang J. Statistical analyses of ballistocardiogram features for cardiac disease diagnosis. J Tsinghua Univ 2015;54:633-637 [FREE Full text]
35. Song Y, Ni H, Zhou X, Zhao W, Wang T. Extracting features for cardiovascular disease classification based on ballistocardiography. Presented at: 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing, 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and its Associated Workshops (UIC-ATC-ScalCom); Aug 10-14, 2015; Beijing. [doi: 10.1109/UIC-ATC-ScalCom-CBDCom-IoP.2015.223]
36. Feng J, Huang W, Jiang J, et al. Non-invasive monitoring of cardiac function through ballistocardiogram: an algorithm integrating short-time Fourier transform and ensemble empirical mode decomposition. Front Physiol 2023;14(1201722):37664434. [doi: 10.3389/fphys.2023.1201722]
37. RICHMAN SM, PRESCOTT R, LITTMANN D. Correlation of the ballistocardiogram with the coronary angiogram and electrocardiogram in ischemic heart disease. Am J Cardiol 1963 Jan;11(1):36-42. [doi: 10.1016/0002-9149(63)90028-6] [Medline: 13982157]
38. Scarborough WR, Penneys R, Thomas CB, Baker BM Jr, Mason RE. The cardiovascular effect of induced controlled anoxemia; a preliminary ballistocardiographic study of normal subjects and a few patients with suspected coronary artery disease. Circulation 1951 Aug;4(2):190-210. [doi: 10.1161/01.cir.4.2.190] [Medline: 14859396]

39. Henderson CB. Ballistocardiograms after cigarette smoking in health and in coronary heart disease. Br Heart J 1953 Jul;15(3):278-286. [doi: 10.1136/hrt.15.3.278] [Medline: 13059215]
40. Davis FW Jr, Scarborough WR, Mason RE, Singewald ML, Baker BM Jr. The effects of exercise and smoking on the electrocardiograms and ballistocardiograms of normal subjects and patients with coronary artery disease. Am Heart J 1953 Oct;46(4):529-542. [doi: 10.1016/0002-8703(53)90064-7] [Medline: 13092040]
41. Theorell T, Blunk D, Wolf S. Emotions and cardiac contractility as reflected in ballistocardiographic recordings. Pavlov J Biol Sci 1974;9(2):65-75. [doi: 10.1007/BF03000526] [Medline: 4220366]
42. Mandelbaum H, Mandelbaum RA. Ballistocardiographic response to nitroglycerin in coronary artery disease. Am Heart J 1952 Aug;44(2):257-260. [doi: 10.1016/0002-8703(52)90150-6] [Medline: 14952446]
43. Jacobs HD. Ballistocardiographic observations. Br Heart J 1954 Jan;16(1):79-86. [doi: 10.1136/hrt.16.1.79] [Medline: 13126361]
44. Trakhtenberg SB. The significance of some noninvasive methods in the objective evaluation of treatment of patients with coronary artery disease. Bibl Cardiol 1976(35):197-201. [Medline: 825099]
45. Zhu Y, Fook VFS, Jianzhong EH. Heart rate estimation from FBG sensors using cepstrum analysis and sensor fusion. Presented at: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Aug 26-30, 2014; Chicago, IL. [doi: 10.1109/EMBC.2014.6944838]
46. Shin JH, Choi BH, Lim YG, Jeong DU, Park KS. Automatic ballistocardiogram (BCG) beat detection using a template matching approach. Presented at: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Aug 20-25, 2008; Vancouver, BC, Canada. [doi: 10.1109/IEMBS.2008.4649363] [Medline: 19162866]
47. Chen X, Zhang H, Li Z, Liu S, Zhou Y. Continuous monitoring of heart rate variability and respiration for the remote diagnosis of chronic obstructive pulmonary disease: prospective observational study. JMIR Mhealth Uhealth 2024 Jul 18;12:e56226. [doi: 10.2196/56226] [Medline: 39024559]
48. Goldberger JJ, Cain ME, Hohnloser SH, et al. American Heart Association/American College of Cardiology Foundation/Heart Rhythm Society Scientific Statement on Noninvasive Risk Stratification Techniques for Identifying Patients at Risk for Sudden Cardiac Death. J Am Coll Cardiol 2008 Sep;52(14):1179-1199. [doi: 10.1016/j.jacc.2008.05.003] [Medline: 18804749]
49. Postolache O, Girao PS, Postolache G. New approach on cardiac autonomic control estimation based on BCG processing. Presented at: 2007 Canadian Conference on Electrical and Computer Engineering; Apr 22-26, 2007; Vancouver, BC, Canada. [doi: 10.1109/CCECE.2007.224]
50. Pinheiro E, Postolache O, Girao P. Pulse arrival time and ballistocardiogram application to blood pressure variability estimation. Presented at: 2009 IEEE International Workshop on Medical Measurements and Applications; May 29-30, 2009; Cetraro, Italy. [doi: 10.1109/MEMEA.2009.5167970]
51. Hoog Antink C, Mai Y, Aalto R, et al. Ballistocardiography can estimate beat-to-beat heart rate accurately at night in patients after vascular intervention. IEEE J Biomed Health Inform 2020 Aug;24(8):2230-2237. [doi: 10.1109/JBHI.2020.2970298] [Medline: 32011272]
52. Shin JH, Hwang SH, Chang MH, Park KS. Heart rate variability analysis using a ballistocardiogram during Valsalva manoeuvre and post exercise. Physiol Meas 2011 Aug;32(8):1239-1264. [doi: 10.1088/0967-3334/32/8/015] [Medline: 21743126]
53. Ye N, Sun YG, Yang D. Noninvasive heart rate variability detection device for fatigue driving detection system. AMM 2013;246-247:194-198 [FREE Full text] [doi: 10.4028/www.scientific.net/AMM.246-247.194]
54. Shen G, Yang M, Zhang B. Ballistocardiogram-based heart rate variation monitoring using unsupervised learning. Adv Transdiscipl Eng 2018:320-329. [doi: 10.3233/978-1-61499-898-3-320]
55. Cui H, Wang Z, Yu B, et al. Statistical analysis of the consistency of HRV analysis using BCG or pulse wave signals. Sensors (Basel) 2022 Mar 21;22(6):2423. [doi: 10.3390/s22062423] [Medline: 35336592]
56. Laurino M, Menicucci D, Gemignani A, Carbonaro N, Tognetti A. Moving auto-correlation window approach for heart rate estimation in Ballistocardiography extracted by mattress-integrated accelerometers. Sensors (Basel) 2020 Sep 22;20(18):5438. [doi: 10.3390/s20185438] [Medline: 32971942]
57. Ganti V, Carek AM, Jung H, et al. Enabling wearable pulse transit time-based blood pressure estimation for medically underserved areas and health equity: comprehensive evaluation study. JMIR Mhealth Uhealth 2021 Aug 2;9(8):e27466. [doi: 10.2196/27466] [Medline: 34338646]
58. Nedogoda SV, Konradi AO, Zvartau NE, et al. Optimization of BP control and arterial elasticity with fixed combination therapy perindopril and amlodipine in patients with arterial hypertension and increased pulse wave velocity. Kardiologiia 2017 Mar;57(3):31-38. [Medline: 28762933]
59. Liu SH, Wu YR, Chen WX, Su CH, Chin CL. Using ballistocardiogram and impedance plethysmogram for minimal contact measurement of blood pressure based on a body weight-fat scale. Sensors (Basel) 2023 Feb 19;23(4):2318. [doi: 10.3390/s23042318] [Medline: 36850917]
60. Kim CS, Carek AM, Mukkamala R, Inan OT, Hahn JO. Ballistocardiogram as proximal timing reference for pulse transit time measurement: potential for cuffless blood pressure monitoring. IEEE Trans Biomed Eng 2015 Nov;62(11):2657-2664. [doi: 10.1109/TBME.2015.2440291] [Medline: 26054058]

61. Farki A, Baradaran Kazemzadeh R, Akhondzadeh Noughabi E. A novel clustering-based algorithm for continuous and noninvasive cuff-less blood pressure estimation. J Healthc Eng 2022;2022:3549238. [doi: 10.1155/2022/3549238] [Medline: 35075386]
62. Huang JJ, Syu HY, Cai ZL, See AR. Development of a long term dynamic blood pressure monitoring system using cuff-less method and pulse transit time. Measurement (Lond) 2018 Aug;124:309-317. [doi: 10.1016/j.measurement.2018.04.047]
63. Winokur ES, He DD, Sodini CG. A wearable vital signs monitor at the ear for continuous heart rate and pulse transit time measurements. Presented at: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 28-1, 2012; San Diego, CA. [doi: 10.1109/EMBC.2012.6346527] [Medline: 23366488]
64. Saran V, Kumar G, Dhawan U, Parchani G. Unsupervised identification of cardiac contraction through ballistocardiography. Presented at: 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES); Dec 3-6, 2018; Sarawak, Malaysia. [doi: 10.1109/IECBES.2018.8626652]
65. Vehkaoja A, Rajala S, Kumpulainen P, Lekkala J. Correlation approach for the detection of the heartbeat intervals using force sensors placed under the bed posts. J Med Eng Technol 2013 Jul;37(5):327-333. [doi: 10.3109/03091902.2013.807523] [Medline: 23795695]
66. Lee H, Cho A, Lee S, Whang M. Vision-based measurement of heart rate from ballistocardiographic head movements using unsupervised clustering. Sensors (Basel) 2019 Jul 24;19(15):3263. [doi: 10.3390/s19153263] [Medline: 31344939]
67. Zeng H, Xu W, Dong B, et al. Beat-to-beat heart rate estimation from MZI-BCG signal based on hierarchical clustering. Presented at: 2021 Opto-Electronics and Communications Conference (OECC); Jul 3-7, 2021; Hong Kong. [doi: 10.1364/OECC.2021.JS3F.13]
68. Kim MS, Ann SY, Ha TK, Yi HD, Lee YJ. 0349 Robust inter-beat interval estimation algorithm using clustering method. Sleep 2022 May 25;45(Supplement_1):A157-A157. [doi: 10.1093/sleep/zsac079.346]
69. Zhao T, Fu X, Zhou Y, Zhan J, Chen K, Li Z. Noncontact monitoring of heart rate variability using a Fiber Optic Sensor. IEEE Internet Things J 2023;10(17):14988-14994. [doi: 10.1109/JIOT.2023.3262634]
70. Gilaberte S, Go´mez-Clapers J, Casanella R, Pallas-Areny R. Heart and respiratory rate detection on a bathroom scale based on the ballistocardiogram and the continuous wavelet transform. Presented at: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology; Aug 31 to Sep 4, 2010; Buenos Aires, Argentina. [doi: 10.1109/iembs.2010.5626866]
71. Alvarado-Serrano C, Luna-Lozano PS, Pallàs-Areny R. An algorithm for beat-to-beat heart rate detection from the BCG based on the continuous spline wavelet transform. Biomed Signal Process Control 2016 May;27:96-102. [doi: 10.1016/j.bspc.2016.02.002]
72. Malik AR, Boger J. Zero-effort ambient heart rate monitoring using ballistocardiography detected through a seat cushion: prototype development and preliminary study. JMIR Rehabil Assist Technol 2021 May 31;8(2):e25996. [doi: 10.2196/25996] [Medline: 34057420]
73. Sadek I, Abdulrazak B. A comparison of three heart rate detection algorithms over ballistocardiogram signals. Biomed Signal Process Control 2021 Sep;70:103017. [doi: 10.1016/j.bspc.2021.103017]
74. Elnaggar I, Hurnanen T, Sandelin J, et al. Multichannel bed based ballistocardiography heart rate estimation using continuous wavelet transforms and autocorrelation. Presented at: 2022 Computing in Cardiology Conference; Sep 4-7, 2022; Tampere, Finland. [doi: 10.22489/CinC.2022.364]
75. Gonzalez-Landaeta R, Casas O, Pallas-Areny R. Heart rate detection from an electronic weighing scale. Presented at: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 22-26, 2007; Lyon, France. [doi: 10.1109/IEMBS.2007.4353791]
76. Jin J, Wang X, Li S, Wu Y. A novel heart rate detection algorithm in ballistocardiogram based on wavelet transform. Presented at: 2009 Second International Workshop on Knowledge Discovery and Data Mining (WKDD); Jan 23-25, 2009; Moscow, Russia.
77. Bruser C, Stadlthanner K, Brauers A, Leonhardt S. Applying machine learning to detect individual heart beats in ballistocardiograms. Presented at: Proc Annu Int Conf IEEE Eng Med Biol Soc (EMBC); Aug 31 to Sep 4, 2010; Buenos Aires, Argentina. [doi: 10.1109/IEMBS.2010.5628077]
78. Brüser C, Stadlthanner K, de Waele S, Leonhardt S. Adaptive beat-to-beat heart rate estimation in ballistocardiograms. IEEE Trans Inf Technol Biomed 2011 Sep;15(5):778-786. [doi: 10.1109/TITB.2011.2128337] [Medline: 21421447]
79. Nagura M, Mitsukura Y, Kishimoto T, Mimura M. An estimation of heart rate variability from ballistocardiogram measured with bed leg sensors. Presented at: 2018 IEEE International Conference on Industrial Technology (ICIT); Feb 20-22, 2018; Lyon, France. [doi: 10.1109/ICIT.2018.8352495]
80. Seo Y, Lee J, Sunarya U, Lee K, Park C. Continuous blood pressure estimation using 1D convolutional neural network and attention mechanism. IEIESPC 2022 Jun 30;11(3):169-173. [doi: 10.5573/IEIESPC.2022.11.3.169]
81. Sakajiri Y, Nakamura H, Ueno A. Non-contact capacitive ballistocardiogram measurements using in-bed fabric-sheet electrode for blood pressure estimation. Presented at: 2019 41st Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); Jul 23-27, 2019; Berlin, Germany. [doi: 10.1109/EMBC.2019.8856770]

82. Jia W, Li Y, Bai Y, Mao ZH, Sun M, Zhao Q. Estimation of heart rate from a chest-worn inertial measurement unit. Presented at: 2015 International Symposium on Bioelectronics and Bioinformatics (ISBB); Oct 14-17, 2015; Beijing, China. [doi: 10.1109/ISBB.2015.7344945]
83. Toreyin H, Javaid AQ, Ashouri H, Ode O, Inan OT. Towards ubiquitous blood pressure monitoring in an armband using pulse transit time. Presented at: 2015 IEEE Biomedical Circuits and Systems Conference (BioCAS); Oct 22-24, 2015; Atlanta, GA, USA. [doi: 10.1109/BioCAS.2015.7348425]
84. Martin SLO, Carek AM, Kim CS, et al. Weighing scale-based pulse transit time is a superior marker of blood pressure than conventional pulse arrival time. Sci Rep 2016 Dec 15;6(39273):39273. [doi: 10.1038/srep39273] [Medline: 27976741]
85. Fierro G, Silveira F, Armentano R. Low group delay signal conditioning for wearable central blood pressure monitoring device. Presented at: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Jul 11-15, 2017; Jeju, South Korea. [doi: 10.1109/EMBC.2017.8037558]
86. Su BY, Enayati M, Ho KC, et al. Monitoring the relative blood pressure using a hydraulic bed sensor system. IEEE Trans Biomed Eng 2019 Mar;66(3):740-748. [doi: 10.1109/TBME.2018.2855639] [Medline: 30010544]
87. Yousefian P, Shin S, Mousavi AS, et al. Physiological association between limb ballistocardiogram and arterial blood pressure waveforms: a mathematical model-based analysis. Sci Rep 2019 Mar 26;9(1):5146. [doi: 10.1038/s41598-019-41537-y] [Medline: 30914687]
88. Carek AM, Jung H, Inan OT. A reflective photoplethysmogram array and channel selection algorithm for weighing scale based blood pressure measurement. IEEE Sensors J 2020;20(7):3849-3858. [doi: 10.1109/JSEN.2019.2960063]
89. Shin S, Mousavi A, Lyle S, et al. Posture-dependent variability in wrist ballistocardiogram-photoplethysmogram pulse transit time: implication to cuff-less blood pressure tracking. IEEE Trans Biomed Eng 2022 Jan;69(1):347-355. [doi: 10.1109/TBME.2021.3094200] [Medline: 34197317]
90. Janjua GMW, Finlay D, Guldenring D, Haq AU, McLaughlin J. Evaluation of pulse transit time for different sensing methodologies of arterial waveforms. IEEE Access 2023;11:33928-33933. [doi: 10.1109/ACCESS.2023.3264291]
91. Steffensen TL, Schjerven FE, Flade HM, et al. Wrist ballistocardiography and invasively recorded blood pressure in healthy volunteers during reclining bike exercise. Front Physiol 2023;14(1189732):1189732. [doi: 10.3389/fphys.2023.1189732] [Medline: 37250120]
92. Lee KJ, Roh J, Cho D, Hyeong J, Kim S. A chair-based unconstrained/nonintrusive cuffless blood pressure monitoring system using a two-channel ballistocardiogram. Sensors (Basel) 2019 Jan 31;19(3):595. [doi: 10.3390/s19030595] [Medline: 30708934]
93. Seok W, Lee KJ, Cho D, Roh J, Kim S. Blood pressure monitoring system using a two-channel ballistocardiogram and convolutional neural networks. Sensors (Basel) 2021 Mar 25;21(7):2303. [doi: 10.3390/s21072303] [Medline: 33806118]
94. Brown HR Jr, Hoffman MJ, de Lalla V Jr. Ballistocardiographic findings in patients with symptoms of angina pectoris. Circulation 1950 Jan;1(1):132-140. [doi: 10.1161/01.cir.1.1.132] [Medline: 15401201]
95. Moss AJ. Ballistocardiographic evaluation of the cardiovascular aging process. Circulation 1961 Mar;23:434-451. [doi: 10.1161/01.cir.23.3.434] [Medline: 13773124]
96. Penneys R, Thomas CB. The relationship between the arterial oxygen saturation and the cardiovascular response to induced anoxemia in normal young adults. Circulation 1950 Mar;1(3):415-425. [doi: 10.1161/01.cir.1.3.415] [Medline: 15405709]
97. Albrecht UV, Mielitz A, Rahman KMA, Kulau U. Identifying Gravity-Related ARTIFACTS on Ballistocardiography Signals by Comparing Weightlessness and Normal Gravity Recordings (ARTIFACTS): Protocol for an Observational Study. JMIR Res Protoc 2024 Sep 26;13:e63306. [doi: 10.2196/63306] [Medline: 39326041]


## Abbreviations

**ACS:** acute coronary syndrome
**AI:** artificial intelligence
**BCG:** ballistocardiogram
**BP:** blood pressure
**CAD:** coronary artery disease
**CHD:** coronary heart disease
**ECG:** electrocardiogram
**HR:** heart rate
**HRV:** heart rate variability
**IHD:** ischemic heart disease
**MI:** myocardial infarction
**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses
**PTT:** pulse transit time

*Edited by KC Wong; submitted 31.10.24; peer-reviewed by FR Sigit Prakoeswa, S Sonawani, SA Ajagbe; revised version received 04.06.25; accepted 04.06.25; published 08.08.25.*

*Please cite as:*
*Maimaiti M, Zhu H, Zeng H*
*Applications of Ballistocardiogram in the Diagnosis of Coronary Heart Disease: Systematic Review*
*JMIR Cardio 2025;9:e68197*
*URL: https://cardio.jmir.org/2025/1/e68197*
*doi:10.2196/68197*

# Physicians' Use of Electronic Health Record Data Elements and Decision Support Tools in Heart Failure Management: User-Centered Cross-Sectional Survey Study

Mohamed S Ali[1], PharmD, SM; Bruna Oewel[2], MDes; Kaitlyn M Greer[1], PhD; Sabah Ganai[1], MPH; Mark W Newman[3], PhD; Kelly Murdoch-Kitt[2], MGD; Scott L Hummel[4,5,6], MD, MS; Michael P Dorsch[1,5], MS, PharmD

[1]College of Pharmacy, University of Michigan, 428 Church Street, Ann Arbor, MI, United States

[2]Penny W. Stamps School of Art & Design, University of Michigan, Ann Arbor, MI, United States

[3]School of Information, University of Michigan, Ann Arbor, MI, United States

[4]Medical School, University of Michigan, Ann Arbor, MI, United States

[5]Frankel Cardiovascular Center, University of Michigan, Ann Arbor, MI, United States

[6]VA Ann Arbor Healthcare System, Ann Arbor, MI, United States

**Corresponding Author:**
Michael P Dorsch, MS, PharmD
College of Pharmacy, University of Michigan, 428 Church Street, Ann Arbor, MI, United States

## Abstract

**Background:** The management of heart failure (HF) requires complex, data-driven decision-making. Although electronic health record (EHR) systems and clinical decision support (CDS) tools can streamline access to essential clinical information, it remains unclear which EHR elements and tools cardiologists and general medicine physicians prioritize when caring for patients with HF.

**Objective:** This study aims to identify these elements and tools to improve the user interface design of future EHR applications.

**Methods:** This study used a user-centered design research approach to understand physician workflows and decision-making needs in HF care. A cross-sectional online survey was administered to 302 physicians, comprising 150 cardiologists (including 15 HF specialists) and 152 general medicine physicians. Respondents reported their use of EHR variables (eg, medication lists, laboratory results, diagnostic tests, problem lists, clinical notes) for decision-making in HF care, as well as their time spent in the EHR before, during, and after patient visits along with their use of predictive models and patient-reported outcome questionnaire. Descriptive analyses, $\chi^2$ tests, and *t* tests were conducted to compare groups, with statistical significance set at $P<.05$.

**Results:** A total of 302 health care providers participated in the survey, nearly evenly split between cardiologists (49.7%, 150/302) and general medicine physicians (50.3%, 152/302). Both groups consistently relied on medication lists, vital signs, laboratory results, diagnostic tests, problem lists, and clinical notes for HF decision-making. Cardiologists placed greater emphasis on diagnostic tests for inpatient HF care (mean [SD] overall frequency, 4.66 [0.50] vs 4.44 [0.64]; $P=.012$) and outpatient HF care (mean [SD] overall frequency, 4.67 [0.55] vs 4.35 [0.71], $P<.001$). In contrast, general medicine physicians relied more on problem lists for inpatient HF care (mean [SD] overall frequency, 4.63 [0.58] vs 4.43 [0.72], $P=.034$), with no significant difference in the outpatient setting ($P>.05$). Both groups underutilized standardized questionnaires and predictive models, with only 20.1% (29/144) of cardiologists and 4.5% (6/133) of general medicine physicians using standardized questionnaires ($P<.001$)

**Conclusions:** Both physician groups depend on medication lists, laboratory results, diagnostic tests, and problem lists. Cardiologists prioritize diagnostic tests, whereas general medicine physicians more often use problem lists. Low use of questionnaires and predictive models highlights the need for better integration of these tools. Future EHR design interface should tailor functionalities to accommodate these differing priorities and optimize HF care.

***(JMIR Cardio 2025;9:e79239)*** doi:10.2196/79239

**KEYWORDS**

heart failure management; electronic health records; clinical decision support; cardiologists; general medicine physicians; workflow integration; health care technology; predictive models; provider satisfaction; user-centered design; user interface design; user experience design; health informatics

## Introduction

### Background

Heart failure (HF) affects more than 64 million people globally [1], placing a significant burden on health care systems [2,3]. In the United States, HF is a leading cause of hospitalization among older adults [4] and is associated with high rates of morbidity and mortality [2,3]. Effective management of HF requires timely and accurate decision-making based on multiple clinical variables, including patient history, laboratory results, imaging, and treatment adherence [5-7]. Ensuring that this critical information is readily available to health care providers in a clear and timely manner is a crucial step in achieving optimal HF management.

An electronic health record (EHR) is an electronic version of a patient's medical history that is maintained by the health care provider over time [8]. It includes all key administrative and clinical data relevant to the patient's care under that provider, such as demographics, progress notes, problem lists, medications, vital signs, past medical history, immunizations, laboratory results, and radiology reports [8,9]. By automating access to information, the EHR has the potential to streamline clinical workflows and make clinical information readily available to health care providers [8,10]. It can also support a range of care-related activities—either directly or indirectly—through interfaces designed for evidence-based decision support [10].

A clinical decision support system (CDSS) delivers timely information, usually at the point of care, to assist health care providers in making informed decisions about a patient's care [11]. When integrated with an EHR, a CDSS can access relevant patient data, highlight key clinical information, and provide tailored recommendations to health care providers [12].

### Knowledge Gap

The first step in designing a new health information technology system for CDSS is to identify the needs of users and define the system's intended functions. A design process includes and revolves around communication with end-users to ascertain their behaviors, motivations, pain points, and needs, as a user-centered design. A user-centered design is an approach for developing applications that incorporates user-centered activities throughout the entire development process [13]. This approach enables end-users to shape the design, enhancing overall usability [14]. In HF, decision-making largely depends on clinical variables and tools, especially when initiating or titrating medications [5-7]. However, identifying the specific EHR variables and tools that cardiologists and general medicine physicians prioritize during HF management remains essential. While prior work has characterized general specialty differences in EHR use [15], the limited uptake of risk prediction tools [16], and patient-reported outcomes in routine practice [17], no study has, for HF specifically, quantified which EHR data elements clinicians deem most important, nor contrasted cardiology versus general medicine priorities across inpatient and outpatient care.

### Objective

Accordingly, this study aimed to determine which EHR information and tools these providers consider most important in their clinical decision-making while caring for patients with HF. Consequently, future CDSSs can be designed and developed accordingly.

## Methods

### Study Design

We used a mixed methods, user-centered approach that incorporated a variety of quantitative and qualitative techniques. This study was a cross-sectional survey of a diverse pool of physicians. We recruited a group of cardiologists, HF specialists, and general medicine physicians through Dynata (Dynata, Shelton, CT, USA), a large data firm that maintains survey participant panels, to answer clinical scenario questions in an online survey. The anonymous online survey was developed using Qualtrics. We stopped the survey once we reached 302 physicians (150 cardiologists and 152 general medicine physicians). Because Dynata recruits participants until a predefined quota is achieved, the response rate information is not available for this study.

### Survey Instrument

The survey was developed and informed based on insights from earlier phases of user-centered research to reduce provider workload, enhance provider decision-making, and improve patient care. This included provider (physician, nurse, pharmacist, or physician assistant) observations, interviews, and prototyping at the University of Michigan, the Veterans Affairs Ann Arbor Healthcare System, St. Joseph Mercy Ann Arbor Hospital, and Henry Ford Hospital and a review of existing literature on EHR-based CDSS for HF management. Specific areas observed included the tools the EHR providers use both inside and outside to care for patients with HF, the pain points encountered when using the EHR system to manage patients with HF, the time required to review a patient in the EHR, and the tools used to streamline workflow in patient management.

The questionnaire (Multimedia Appendix 1) consisted of several domains, including the provider's clinical role, whether they treat patients with HF, the percentage of time they spend on HF care, and the types of patients they manage (inpatients, outpatients, or both). It also gathered information on years of experience using EHRs, the amount of time spent interacting with the EHR before, during, and after patient visits, the number of software applications and predictive tools used daily, and the most frequently used EHR vendor. Most importantly, the survey included a Likert-scale question assessing how frequently providers rely on a list of clinical information for HF treatment decisions in both inpatient and outpatient settings. Participants were recruited from a national online physician panel (Dynata). The survey did not implement stratified sampling by geographic region, institutional type (eg, academic vs community), or practice ownership. To enable planned comparisons, the survey used a quota by specialty (cardiology vs general medicine); otherwise, enrollment was consecutive until the target sample

size was reached. The study reported practice setting (inpatient, outpatient, or both), EHR vendor, and years of EHR use to characterize sample diversity. The survey was validated through an independent expert review by HF cardiologists, HF clinical pharmacists, and survey methodologists; we incorporated their feedback via iterative revisions to improve clarity, relevance, and completeness.

### Statistical Analysis

Descriptive statistics were used to summarize the baseline characteristics of the participants, which included their clinical roles, the time providers spent caring for patients with HF, the EHR systems used, and the duration of EHR usage before, during, and after patient visits. Frequencies and percentages were calculated for each category.

Our primary variable of interest was the clinical variables that providers used in clinical decision-making for HF and how these were prioritized. Responses to survey questions related to our primary variable of interest were captured using a 5-point Likert scale. The scale ranged from 1 to 5, where 1=never, 2=rarely, 3=sometimes, 4=often, and 5=always. Providers were asked to indicate how frequently they used that information in their decision-making process for each type of Likert-scaled clinical information (eg, diagnostic tests, laboratory results, and clinical notes). The responses were numerically coded according to the Likert scale, and the mean score for each group (cardiologists and general medicine physicians) was calculated by averaging the numerical responses. The SD was also calculated to indicate the variability in responses within each variable. Our secondary variable of interest was the time spent on EHRs before, during, and after patient visits, as well as the number of software systems used daily by both cardiologists and general medicine physicians.

The $\chi^2$ tests were used to assess differences in categorical variables for comparison between provider groups (eg, cardiologists vs general medicine physicians). The *t* tests were used for continuous variables, with statistical significance set at $P<.05$. R programming (version 4.4.2; R Foundation for Statistical Computing) and SAS (version 9.4; SAS Institute Inc.) were used to conduct these analyses and generate the figures.

### Ethical Considerations

This study was deemed exempt by the University of Michigan Institutional Review Board because it was not found to constitute human subject research. All participants provided electronic informed consent within Qualtrics before beginning the survey. Recruitment and compensation were managed by Dynata; participants received panel-standard incentives, and the study team did not access personal contact information or payment details.

## Results

### Participants

A total of 302 health care providers participated in the survey with cardiologists representing 49.7% (150/302) of the population and general medicine physicians representing 50.3% (152/302) of the population. Of the general medicine group, 50% (76/152) were family medicine physicians and 50% (76/152) were internal medicine physicians. Among cardiologists, 10% (15/150) specialized in HF. In terms of time spent managing HF, 42.7% (129/302) of respondents spent 1% - 24% of their time, 37.4% (113/302) spent 25% - 49%, 10.3% (31/302) spent 50% - 74%, and 9.6% (29/302) spent 75% - 100%. Most (66.6%, 197/302) provided both inpatient and outpatient care; 27% (80/302) cared only for outpatients, and 6.4% (19/302) only for inpatients.

### EHR Use Across the Visit Workflow

EHR usage varied across the clinical workflow. Before visits, 58% of cardiologists (87/150) and 64% of general medicine physicians (97/152) reported spending 3 - 10 minutes in the EHR, whereas during visits, usage was typically between 1 and 5 minutes. After visits, 61% of the respondents (184/302) reported 3 - 10 minutes of use. Overall, 92.4% of participants (279/302) had over 5 years of experience with EHRs. Detailed distributions are provided in Table 1. The time spent reviewing EHR before, during, and after patients' visits is shown in Table 2.

**Table .** Baseline provider demographics.

| Variable | Value, n (%; n=302) |
|---|---|
| Cardiologist | 150 (49.7) |
| HF Cardiologist | 15 (10.0) |
| General Medicine | 152 (50.3) |
| Family Medicine | 76 (50.0) |
| Internal Medicine | 76 (50.0) |
| Percent of time the provider spent caring for patients with heart failure | |
| 1% - 24% | 129 (42.7) |
| 25% - 49% | 113 (37.4) |
| 50% - 74% | 31 (10.3) |
| 75% - 100% | 29 (9.6) |
| Provider care setting | |
| Inpatient only | 19 (6.4) |
| Outpatient only | 80 (27.0) |
| Both | 197 (66.6) |
| Length of time the provider has been using an EHR[a] | |
| 0 - 2 years | 1 (0.3) |
| 3 - 5 years | 19 (6.3) |
| >5 years | 279 (92.4) |
| Unknown | 3 (1.0) |
| EHR company used most frequently | |
| Allscripts | 29 (9.6) |
| Cerner | 39 (13.0) |
| Epic | 142 (47.2) |
| MEDITECH | 18 (6.0) |
| Other | 73 (24.2) |
| Number of different software systems used daily | |
| 1 | 86 (28.6) |
| 2 | 105 (34.9) |
| 3 | 73 (24.3) |
| 4 | 20 (6.6) |
| 5 or more | 17 (5.6) |

[a]EHR: electronic health record.

**Table .** EHR[a] usage before, during, and after patient visits.

| Time spent in the EHR | n (%) |
|---|---|
| Prior to a patient visit | |
| 0 minutes | 5 (1.7) |
| 1-3 minutes | 62 (20.7) |
| 3-5 minutes | 95 (31.7) |
| 5-10 minutes | 88 (29.3) |
| >10 minutes | 50 (16.7) |
| While in a patient visit | |
| 0 minutes | 33 (11) |
| 1-3 minutes | 82 (27.2) |
| 3-5 minutes | 79 (26.2) |
| 5-10 minutes | 65 (21.6) |
| >10 minutes | 42 (14) |
| After a patient visit | |
| 0 minutes | 12 (4) |
| 1-3 minutes | 52 (17.2) |
| 3-5 minutes | 92 (30.5) |
| 5-10 minutes | 92 (30.5) |
| >10 minutes | 54 (17.9) |

[a]EHR: electronic health record.

## Information Elements Used for HF Decisions

Both cardiologists and general medicine physicians, across inpatient and outpatient care, consistently used medication lists, vital signs, laboratory results, diagnostic tests, problem lists, clinical notes, medical history, and allergies in their clinical practice.

Cardiologists more frequently relied on diagnostic tests for inpatient HF treatment decisions than general medicine physicians (mean [SD] overall frequency, 4.66 [0.50] vs 4.44 [0.64]; *P*=.012). However, general medicine physicians relied on problem lists more than cardiologists (mean [SD] overall frequency, 4.63 [0.58] vs 4.43 [0.72]; *P*=.034). In contrast, there was no significant difference between both in the use of other variables (eg, medication lists [*P*=.098], vital signs [*P*=.420], and laboratory results [*P*=.244]). Figures 1 and 2 demonstrate the clinical variables that cardiologists and general medicine physicians relied on for clinical decision-making in patients with HF.

**Figure 1.** Frequency providers used clinical information in their decision-making process for inpatients. The figure illustrates the frequency of reviewing EHR data elements by cardiologists and general medicine physicians in an inpatient setting. Data elements are ordered by overall frequency of use and shown as mean (SD). EHR: electronic health record; GP: general medicine physician.

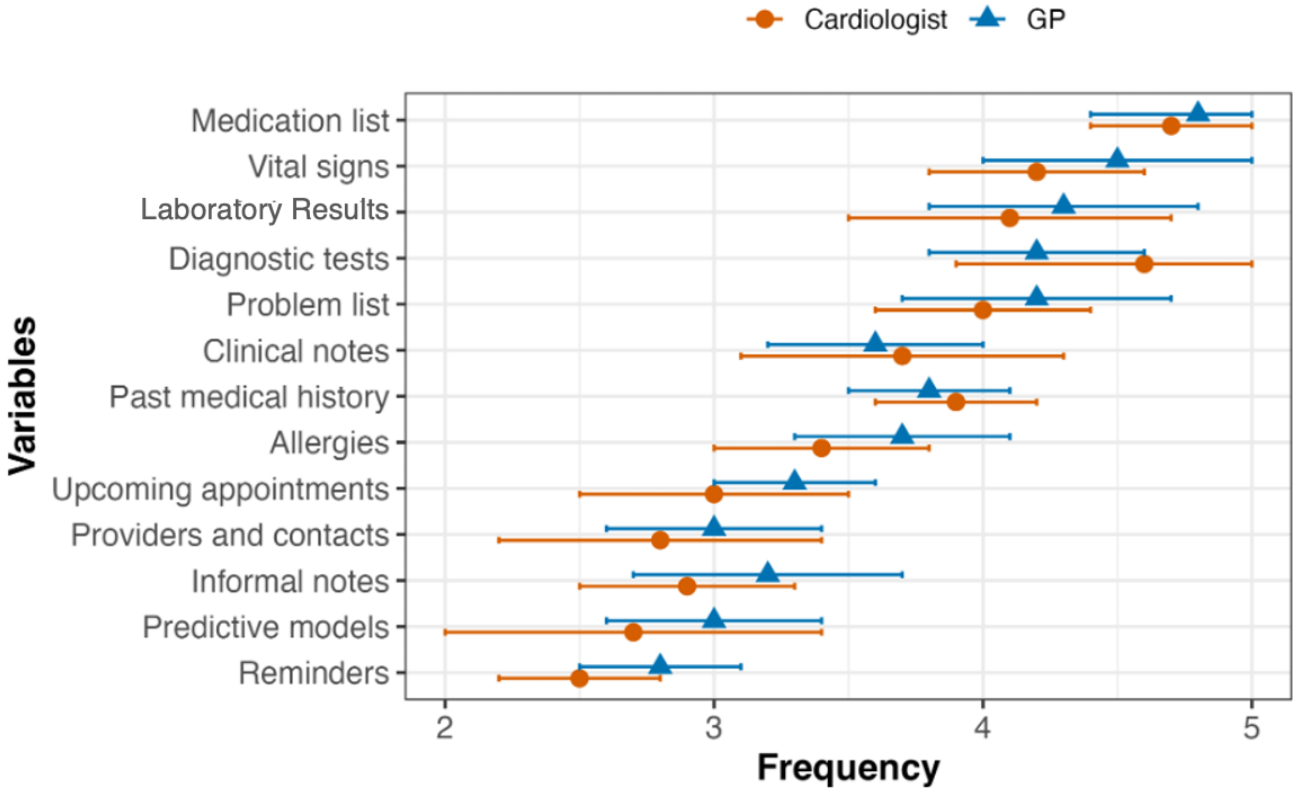
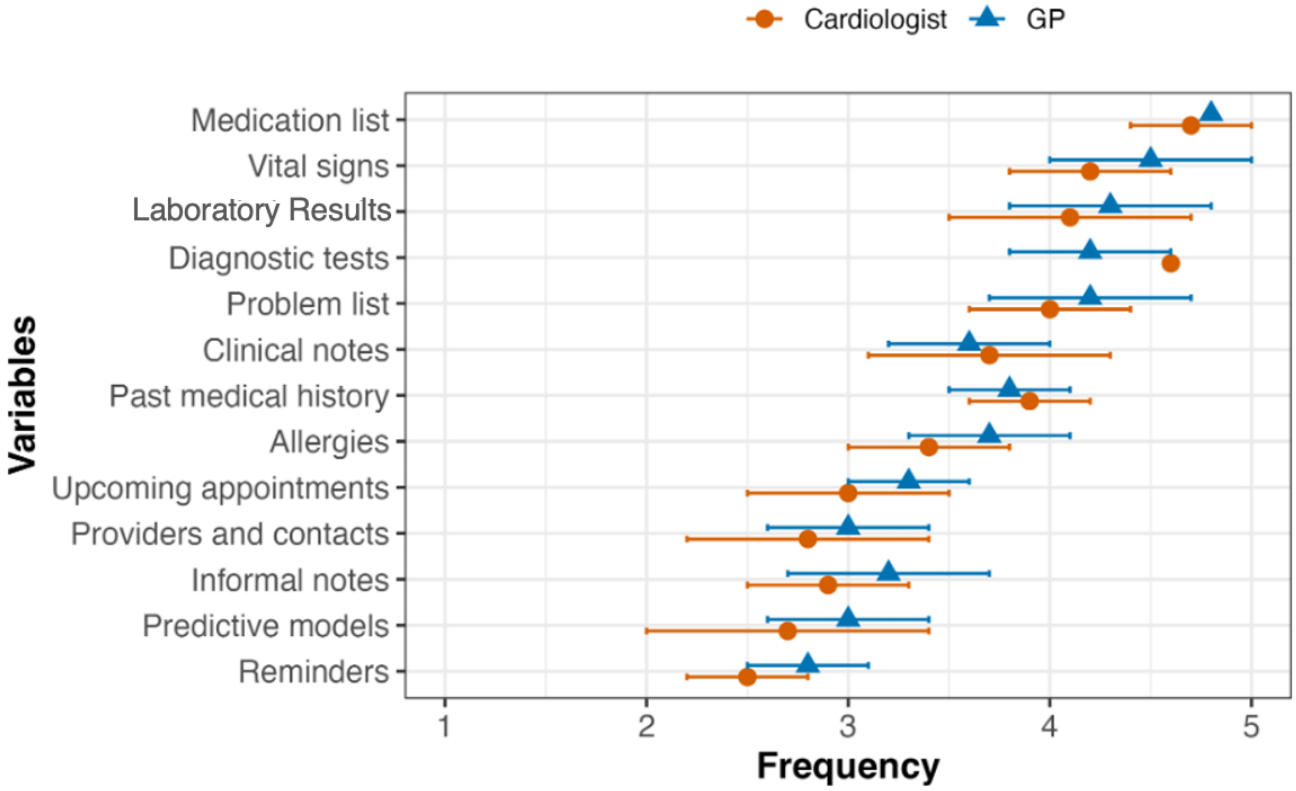



**Figure 2.** Frequency providers used clinical information in their decision-making process for outpatients. The figure illustrates the frequency of reviewing EHR data elements by cardiologists and general medicine physicians in an outpatient setting. Data elements are ordered by overall frequency of use and shown as mean (SD). EHR: electronic health record; GP: general medicine physician.




Cardiologists also relied more on diagnostic tests for outpatient treatment decisions than general medicine physicians (mean [SD] overall frequency, 4.67 [0.55] vs 4.35 [0.71]; *P*<.001). Additionally, general medicine physicians reviewed clinical notes less frequently than cardiologists (mean [SD] overall frequency, 4.50 [0.67] vs 4.65 [0.64]; *P*=.042).

### Use of Standardized Questionnaires and Predictive Models

Both provider groups underutilized standardized questionnaires and predictive models. Only 20.1% (29/144) of cardiologists and 4.5% (6/133) of general medicine physicians reported using standardized questionnaires, resulting in a significant difference (*P*<.001). Predictive model usage was similarly low, at 23.1% (33/144) among cardiologists and 27.3% (36/133) among general medicine physicians, with no significant difference between the groups (*P*=.423). The denominators (144 cardiologists and 133 general medicine physicians) reflect the subset of respondents who completed these specific survey questions; 8.3% of participants (25/302) skipped these questions.

### Time in EHR by Specialty

Cardiologists spent significantly more time reviewing the EHR before patient visits than general medicine physicians. Among cardiologists, 31.8% (47/150) spent 5 - 10 minutes reviewing the EHR before patient visits, whereas 36.8% (56/152) of general medicine physicians spent only 3 - 5 minutes (*P*=.035). However, there were no statistically significant differences in EHR usage during or after visits between the two groups (*P*=.247 and *P*=.170, respectively). Table 3 illustrates this in more detail.

### Number of Software Systems Used Daily

Cardiologists were more inclined to use multiple software systems, with 31.5% (47/150) using three different systems, while 40.1% (61/152) of general medicine physicians used only 2 (*P*=.001). Table 3 shows the number of different software systems used daily.

**Table .** Time spent in the EHR[a] and number of software systems used daily by specialty.

| Variable | Cardiologist (n=150) | General medicine physician (n=152) | *P* value |
|---|---|---|---|
| Time spent in the EHR prior to a patient visit | | | 0.035 |
| 0 minutes | 3 (2) | 2 (1.3) | |
| 1-3 minutes | 26 (17.6) | 36 (23.7) | |
| 3-5 minutes | 39 (26.4) | 56 (36.8) | |
| 5-10 minutes | 47 (31.8) | 41 (27) | |
| >10 minutes | 33 (22.3) | 17 (11.2) | |
| Time spent in the EHR while in a patient visit | | | 0.247 |
| 0 minutes | 14 (9.4) | 19 (12.5) | |
| 1-3 minutes | 46 (30.9) | 36 (23.7) | |
| 3-5 minutes | 39 (26.2) | 40 (26.3) | |
| 5-10 minutes | 26 (17.5) | 39 (25.7) | |
| >10 minutes | 24 (16.1) | 18 (11.8) | |
| Time spent in the EHR after a patient visit | | | 0.17 |
| 0 minutes | 4 (2.7) | 8 (5.3) | |
| 1-3 minutes | 27 (18) | 25 (16.5) | |
| 3-5 minutes | 49 (32.7) | 43 (28.3) | |
| 5-10 minutes | 38 (25.3) | 54 (35.5) | |
| >10 minutes | 32 (21.3) | 22 (14.5) | |
| The number of different software systems used on a daily basis | | | 0.001 |
| 1 | 33 (22.2) | 53 (34.9) | |
| 2 | 44 (29.5) | 61 (40.1) | |
| 3 | 47 (31.5) | 26 (17.1) | |
| 4 | 14 (9.4) | 6 (3.9) | |
| 5 or more | 11 (7.4) | 6 (3.9) | |

[a]EHR: electronic health record.

## Discussion

### Principal Findings

Both cardiologists and general medicine physicians rely heavily on specific EHR data, including medication lists, vital signs, laboratory results, diagnostic tests, problem lists, and clinical notes, when managing HF. Accordingly, new EHR interfaces should be designed to ensure these essential elements are presented in a clear, accessible format that supports timely clinical decision-making. In contrast to previous research suggesting that specialists and primary care physicians regard a patient's clinical history—chief complaint, history of present illness, and past medical history—as among the most critical EHR sections for HF evaluation [18], our study found that both provider groups used past medical history less frequently.

### Comparison With Prior Work

While both provider groups relied on similar categories of information, our survey highlighted notable differences in how cardiologists and general medicine physicians prioritize EHR data and tools. Specifically, cardiologists placed greater emphasis on diagnostic tests, whereas general medicine physicians more frequently used the problem list. These findings are consistent with a multi-specialty survey in which 50% of specialists ranked imaging data among their top five information needs, compared with only 27% of primary care physicians [18]. Conversely, 61% of primary care physicians ranked the problem list in their top five, as opposed to just 27% of specialists [18]. This focus on the problem list aligns with the broader, long-term management responsibilities typically associated with primary care physicians. Indeed, one large study showed primary care physicians entering over 80% of all problem list items, whereas specialists contributed relatively few [19]. These differences highlight the importance of designing EHR interfaces that accommodate the distinct workflows and data needs of both specialists and general medicine physicians.

Differences also emerged in the use of software systems. Our survey suggests that cardiologists tend to use a greater number of software systems during HF care. For example, a cardiologist might navigate the primary EHR for notes and orders, a separate cardiology picture archiving and communication system for imaging [20], and device-specific platforms for pacemaker and *International Classification of Diseases* data [21]. General medicine physicians, by comparison, usually work within a single EHR ecosystem for most tasks. This describes the "network of systems" approach in specialty care: a one-size-fits-all EHR often fails to meet all specialty needs, leading many specialists to adopt "best-of-breed" solutions (multiple integrated systems tailored to their domain) [18]. In contrast, general medicine physicians may engage more with general CDSS alerts or chronic disease management prompts embedded in the EHR (eg, health maintenance reminders, drug-interaction alerts). Prior research has noted that primary care physicians place higher value on medication-related information and may be more receptive to certain decision support tied to the problem list or medication list [18].

Our survey indicates that cardiologists spend significantly more time reviewing the EHR prior to patient visits than general medicine physicians. This confirms that certain medical subspecialties experience high EHR workloads. In a large cross-specialty analysis, infectious disease, endocrinology, and nephrology were among the top specialties for total EHR time, on par with or exceeding primary care [22]. These fields, much like cardiology, manage complicated patients with multiple comorbidities and large volumes of data, which naturally translates into more time spent reviewing results, notes, and orders in the EHR [22]. Enhanced EHR design and workflow customization tailored to specialty needs may be beneficial. For example, cardiology-specific dashboards that compile recent cardiac test results, or smarter integration of hospital records and consult notes, could streamline pre-visit preparation for cardiologists. Likewise, adopting team-based planning elements in cardiology clinics (where appropriate) might offload routine data gathering from physicians. The goal would be to reduce the unnecessary time clinicians spend clicking and searching in the chart, thereby improving efficiency without sacrificing thoroughness.

Our survey revealed the underutilization of standardized questionnaires (eg, patient-reported outcome measures) and predictive risk models by both cardiologists and general medicine physicians in routine HF care. Despite the proliferation of risk prediction tools and symptom questionnaires for HF, their adoption in everyday practice remains low. Dozens of HF risk models (for outcomes such as mortality or readmission) have been published, but clinicians rarely incorporate these models at the point of care [23]. One review noted that there is no clear guideline consensus on which risk score to use, and in a large European HF registry, fewer than 1% of patients had any prognostic risk score documented in their medical record [23]. Clinicians often cite multiple barriers: predictive models developed in research may lack perceived reliability for individual patients, can be too complex or inconvenient, and may not readily fit into clinical workflows [23]. In HF, physicians may also feel that risk stratification adds little to their clinical judgment—for instance, some may perceive that patients with HF are high risk by default, so a calculated risk score might not change management [24]. This skepticism, combined with the absence of strong guideline recommendations for specific models, leads to very limited use of risk calculators at the bedside. Similarly, standardized questionnaires such as the Kansas City Cardiomyopathy Questionnaire (KCCQ) or other quality-of-life instruments are infrequently used by busy clinicians despite their proven value in research settings. Major HF guidelines encourage the assessment of patient-reported health status using tools like the KCCQ to capture symptoms and quality of life. In practice, however, the routine use of KCCQ and other patient-reported outcome surveys is rare, and these instruments are often reserved for clinical trials or specialized programs [25].

### Implications for EHR/CDSS Design

Aligned with our findings, CDSS changes should mirror observed use patterns: as cardiologists prioritized diagnostic tests, spent more pre-visit EHR time, and used more systems, the CDSS should surface diagnostic test summaries early in the

workflow and aggregate test information to reduce pre-visit review and system switching; because general medicine physicians relied more on the problem list (inpatient), the CDSS should orient guidance around the problem list with clear access from inpatient views. Given that both groups consistently used medication lists, vitals, laboratory results, diagnostic tests, problem lists, and notes, decision cues should be placed adjacent to these high-traffic elements rather than in separate modules. Finally, the low reported use of standardized questionnaires and predictive models suggests placing prompts and access points for these tools on the same screens clinicians already frequent, instead of standalone locations.

### Limitations

This study has several limitations. First, the cross-sectional survey design captures only a snapshot of providers' EHR usage patterns and priorities, limiting the ability to infer causal relationships or temporal changes. Second, data were self-reported, introducing the potential for recall and social desirability biases. Third, the survey was administered through a single vendor (Dynata), which may limit generalizability if respondents are not fully representative of all cardiologists and general medicine physicians. Fourth, patient outcomes or clinical effectiveness measures associated with EHR usage were not assessed, precluding direct links between specific EHR practices and improvements in HF care. Fifth, although the sample included providers from diverse health care systems, differences in EHR functionalities and vendors across institutions may influence how respondents interact with and prioritize EHR data. Sixth, the survey did not stratify by geographic region, institutional type, or practice setting; thus, the panel-derived sample may not perfectly reflect the national distribution of HF specialists and general medicine physicians. While specialty quotas supported planned comparisons, residual sampling bias remains possible. Seventh, although Likert scale data are ordinal, responses were treated as continuous for calculating means and conducting *t* tests. This approach, commonly used in health informatics and social science research, assumes equal intervals between response categories, which may not fully capture participants' subjective perceptions. Results should therefore be interpreted with this methodological consideration in mind. Finally, because participants were recruited from a national US physician panel and the study team is based in the United States, the findings primarily reflect clinical practices and EHR systems in the United States. Therefore, generalizability to other countries with different health system structures and EHR implementations may be limited.

### Conclusions

Cardiologists and general medicine physicians depend on medication lists, vital signs, laboratory results, diagnostic tests, problem lists, and clinical notes to manage HF. Cardiologists place greater emphasis on diagnostic tests, spend more pre-visit EHR time, and use more software systems, whereas general medicine physicians rely more on problem lists for inpatient care. Both groups underutilize standardized questionnaires and predictive models. Tailoring the interface design of the EHR and CDSS tools to these specialty-specific needs could streamline workflows and improve HF management.

## Acknowledgments

MD is supported by R18 HS026874 from the Agency for Health Research and Quality, R33 HL155498 from the NIH/NHLBI, and the American Heart Association Health IT Research Network. Dr. Hummel is supported by a grant from Veterans Affairs, CARA-009- ‐ 16F9050.

## Funding

Financial support for the study was provided by the University of Michigan.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey questions.
[DOCX File, 21 KB - cardio_v9i1e79239_app1.docx ]

## References

1. Roger VL. Epidemiology of heart failure: a contemporary perspective. Circ Res 2021 May 14;128(10):1421-1434. [doi: 10.1161/CIRCRESAHA.121.318172] [Medline: 33983838]
2. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res 2023 Jan 18;118(17):3272-3287. [doi: 10.1093/cvr/cvac013] [Medline: 35150240]
3. Shahim B, Kapelios CJ, Savarese G, Lund LH. Global public health burden of heart failure: an updated review. Card Fail Rev 2023;9:e11. [doi: 10.15420/cfr.2023.05] [Medline: 37547123]
4. Butrous H, Hummel SL. Heart failure in older adults. Can J Cardiol 2016 Sep;32(9):1140-1147. [doi: 10.1016/j.cjca.2016.05.005] [Medline: 27476982]

5. Heidenreich PA, Bozkurt B, Aguilar D, et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. Circulation 2022 May 3;145(18):e895-e1032. [doi: 10.1161/CIR.0000000000001063] [Medline: 35363499]
6. Allen LA, Stevenson LW, Grady KL, et al. Decision making in advanced heart failure: a scientific statement from the American Heart Association. Circulation 2012 Apr 17;125(15):1928-1952. [doi: 10.1161/CIR.0b013e31824f2173] [Medline: 22392529]
7. Sapna F, Raveena F, Chandio M, et al. Advancements in heart failure management: a comprehensive narrative review of emerging therapies. Cureus 2023 Oct;15(10):e46486. [doi: 10.7759/cureus.46486] [Medline: 37927716]
8. Honavar SG. Electronic medical records – the good, the bad and the ugly. Indian J Ophthalmol 2020;68(3):417-418. [doi: 10.4103/ijo.IJO_278_20]
9. Wang W, Ferrari D, Haddon-Hill G, Curcin V. Electronic health records as source of research data. In: Machine Learning for Brain Disorders 2023. [doi: 10.1007/978-1-0716-3195-9_11]
10. Kao DP, Trinkley KE, Lin CT. Heart failure management innovation enabled by electronic health records. JACC Heart Fail 2020 Mar;8(3):223-233. [doi: 10.1016/j.jchf.2019.09.008] [Medline: 31926853]
11. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3(1):17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
12. Chen Z, Liang N, Zhang H, et al. Harnessing the power of clinical decision support systems: challenges and opportunities. Open Heart 2023 Nov 28;10(2):e002432. [doi: 10.1136/openhrt-2023-002432] [Medline: 38016787]
13. Newcomer NL, Lindahl D, Martin K. User centered design. In: Encyclopedia of Database Systems 2009.
14. Lee S. Usability testing for developing effective interactive multimedia software: Concepts, dimensions, and procedures. J Educ Technol Soc 1999 Apr 1;2 [FREE Full text]
15. Sinha S, Holmgren AJ, Hong JC, Rotenstein LS. Ctrl-C: a cross-sectional study of the electronic health record usage patterns of US oncology clinicians. JNCI Cancer Spectr 2023 Aug 31;7(5):pkad066. [doi: 10.1093/jncics/pkad066] [Medline: 37688578]
16. Sharma V, Ali I, van der Veer S, Martin G, Ainsworth J, Augustine T. Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. BMJ Health Care Inform 2021 Feb;28(1):e100253. [doi: 10.1136/bmjhci-2020-100253] [Medline: 33608259]
17. Biliavska I, Lenguerrand E, Tobias JH, Hamann PDH. How and who? Examining real-world evidence of engagement and use of an electronic patient-reported outcome smartphone application in routine clinical care for patients with inflammatory arthritides. ACR Open Rheumatol 2025 May;7(5):e70049. [doi: 10.1002/acr2.70049] [Medline: 40326863]
18. Redd TK, Doberne JW, Lattin D, et al. Variability in electronic health record usage and perceptions among specialty vs. primary care physicians. AMIA Annu Symp Proc 2015;2015:2053-2062. [Medline: 26958305]
19. Wright A, Feblowitz J, Maloney FL, Henkin S, Bates DW. Use of an electronic problem list by primary care providers and specialists. J Gen Intern Med 2012 Aug;27(8):968-973. [doi: 10.1007/s11606-012-2033-5] [Medline: 22426706]
20. Strickland NH. PACS (picture archiving and communication systems): filmless radiology. Arch Dis Child 2000 Jul;83(1):82-86. [doi: 10.1136/adc.83.1.82] [Medline: 10869010]
21. Burri H, Senouf D. Remote monitoring and follow-up of pacemakers and implantable cardioverter defibrillators. Europace 2009 Jun;11(6):701-709. [doi: 10.1093/europace/eup110] [Medline: 19470595]
22. Holmgren AJ, Sinsky CA, Rotenstein L, Apathy NC. National comparison of ambulatory physician electronic health record use across specialties. J Gen Intern Med 2024 Nov;39(14):2868-2870. [doi: 10.1007/s11606-024-08930-4] [Medline: 38980460]
23. Di Tanna GL, Wirtz H, Burrows KL, Globe G. Evaluating risk prediction models for adults with heart failure: a systematic literature review. PLoS ONE 2020;15(1):e0224135. [doi: 10.1371/journal.pone.0224135] [Medline: 31940350]
24. Yap J. Risk stratification in heart failure: existing challenges and potential promise. Int J Cardiol 2020 Aug 15;313:97-98. [doi: 10.1016/j.ijcard.2020.04.042] [Medline: 32320780]
25. Brown-Johnson C, Calma J, Amano A, et al. Evaluating the implementation of patient-reported outcomes in heart failure clinic: a qualitative assessment. Circ: Cardiovascular Quality and Outcomes 2023 May;16(5). [doi: 10.1161/CIRCOUTCOMES.122.009677]


## Abbreviations

**CDSS:** clinical decision support system
**EHR:** electronic health record
**HF:** heart failure
**KCCQ:** Kansas City Cardiomyopathy Questionnaire

*Edited by A Coristine; submitted 17.06.25; peer-reviewed by K Hnid, O Ogorry; revised version received 21.10.25; accepted 21.10.25; published 14.11.25.*

*Please cite as:*
*Ali MS, Oewel B, Greer KM, Ganai S, Newman MW, Murdoch-Kitt K, Hummel SL, Dorsch MP*
*Physicians' Use of Electronic Health Record Data Elements and Decision Support Tools in Heart Failure Management: User-Centered Cross-Sectional Survey Study*
*JMIR Cardio 2025;9:e79239*
*URL: https://cardio.jmir.org/2025/1/e79239*
*doi:10.2196/79239*

# A Web-Based Tool to Perform a Values Clarification for Stroke Prevention in Patients With Atrial Fibrillation: Design and Preliminary Testing Study

Michael P Dorsch[1], MS, PharmD; Allen J Flynn[2], PharmD, PhD; Kaitlyn M Greer[1], PhD; Sabah Ganai[1], MPH; Geoffrey D Barnes[3], MSc, MD; Brian Zikmund-Fisher[4], PhD

[1]College of Pharmacy, University of Michigan, 248 Church Street, Ann Arbor, MI, United States

[2]Department of Learning Health Sciences, Medical School, University of Michigan, Ann Arbor, MI, United States

[3]Division of Cardiovascular Medicine, Medical School, University of Michigan, Ann Arbor, MI, United States

[4]School of Public Health, University of Michigan, Ann Arbor, MI, United States

**Corresponding Author:**
Michael P Dorsch, MS, PharmD
College of Pharmacy, University of Michigan, 248 Church Street, Ann Arbor, MI, United States

## Abstract

**Background:** Atrial fibrillation (AF) is associated with an increased risk of stroke. Oral anticoagulation (OAC) is used for stroke prevention in AF, but it also increases bleeding risk. Clinical guidelines do not definitively recommend for or against OAC for patients with borderline stroke risk. Decision-making may benefit from values clarification exercises to communicate risk trade-offs.

**Objective:** This study aimed to evaluate if a visual with a values clarification alters the understanding of the trade-offs of anticoagulation in AF.

**Methods:** Participants aged 45 - 64 years were recruited across the United States via an online survey. While answering the survey, they were asked to imagine they were newly diagnosed with AF with a $CHA_2DS_2$-VASc (congestive heart failure; hypertension; age ≥75 years [doubled]; type 2 diabetes; previous stroke, transient ischemic attack, or thromboembolism [doubled]; vascular disease; age 65 to 75 years; and sex category) score of 1 for men and 2 for women. Eligibility criteria included no diagnosis of AF and no prior OAC use. Participants were randomized to one of three conditions: (1) standard text-based information only (n=255), (2) visual aids showing stroke-risk probabilities (n=218), or (3) visual aids plus a values clarification exercise (visual+VC; n=200). Participants were subrandomized within the 2 visual-based groups to view either a gauge display or an icon array representing stroke risk. All participants read a hypothetical scenario of being newly diagnosed with AF and hypertension. The primary outcome was decision confidence as measured by the SURE (Sure of Myself; Understand Information; Risk-Benefit Ratio; Encouragement) test. Secondary measures included participants' perceived stroke risk reduction, worry about stroke or bleeding, and likelihood to choose OAC.

**Results:** A total of 673 participants completed the survey. The overall SURE test was 61.2% (156/255) for the standard, 66.5% (145/218) for the visual, and 67% (134/200) for the visual+VC group (visual vs standard *P*=.23; visual+VC vs standard *P*=.20). Participants were less likely to choose OAC in the visual groups (standard: mean 58.3, SD 30; visual: mean 51.4, SD 32; visual+VC: 51.9, SD 28; *P*=.03). Participants felt the reduction in stroke risk from an OAC was less in the visual groups (standard: mean 63.8, SD 22; visual: mean 54.2, SD 28; visual+VC: mean 58.6, SD 25; *P*<.001). Visualization methods (gauge vs icon array) showed no significant differences in overall SURE test results. Participants were less likely to choose OAC and perceived a smaller stroke risk reduction with gauge than icon array (OAC choice: gauge 48.8, icon array 55.4; *P*=.03; stroke risk reduction: gauge 52.1, icon array 60.4; *P*=.001).

**Conclusions:** Visual aids can modestly affect decision confidence and perceptions regarding the benefits of OAC but do not significantly alter decision certainty in a scenario where the guidelines do not recommend for or against OAC. Future work should determine the role of a gauge versus icon array visual for decision-making in stroke prevention in AF.

***(JMIR Cardio 2025;9:e67956)*** doi:10.2196/67956

**KEYWORDS**

digital health; atrial fibrillation; stroke prevention; shared decision-making; values clarification

## Introduction

Risk stratification and shared decision-making are essential in stroke prevention in atrial fibrillation (SPAF). In a wide variety of patients with AF, anticoagulation reduces the risk of ischemic stroke by 65% with a relative 2-fold increase in major extracranial bleeding compared to placebo [1-3]. Yet, medication responses vary across patients. Personalized risks and benefits are available to clinicians via the $CHA_2DS_2$-VASc (congestive heart failure; hypertension; age ≥75 years [doubled]; type 2 diabetes; previous stroke, transient ischemic attack, or thromboembolism [doubled]; vascular disease; age 65 to 75 years; and sex category) and HAS-BLED (hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile international normalized ratio, elderly [>65 years], drugs/alcohol concomitantly) risk scoring systems, representing the risk of stroke and bleeding in AF [4-6]. These tools can provide a tailored estimate of a patient's benefit and risk of anticoagulation in AF.

Many current AF-shared decision-making tools use visual tools such as icon arrays to display the percent risk of stroke ($CHA_2DS_2$-VASc) and risk of bleed (HAS-BLED). While such tools help convey probabilities to patients [7], such probability-focused communications do not visually distinguish between different outcomes. This is a problem because it may lead patients and clinicians to give similar weight to these outcomes even though the medical complications of a stroke are far greater than the medical complications of a bleed. AF guidelines indicate that for the majority of patients where anticoagulation is recommended ($CHA_2DS_2$-VASc ≥2), the HAS-BLED is best used to remove or treat risk factors for bleeding (eg, stop concomitant aspirin or nonsteroidal anti-inflammatory drugs and treat hypertension) rather than to determine if anticoagulation should or should not be given.

One approach to encouraging more thoughtful consideration of the different possible outcomes of AF is using values clarification exercises [3]. Values clarification exercises are structured activities that encourage people to consider how much subjective weight they place on different possible outcomes [8-10]. For many years, developers of patient decision aids have encouraged the inclusion of values clarification exercises in such tools to increase the alignment of medical decisions with patient preferences. However, there is limited evidence on the comparative effectiveness of these different formats in the context of oral anticoagulation (OAC) decision-making in AF.

We report the results of a multistep design and evaluation process to explore the potential for integrating values clarification exercise–derived patient values into presentations of the risks and benefits of anticoagulant therapy. We based our work on the Ottawa Decision Support Framework (ODSF), an evidence-based midrange theory guiding patients' health decisions [11,12]. The framework is based on concepts from psychology, decision analysis, and decision conflict to evaluate the quality of outcomes in providing decision support. In this project, we engaged patients and providers in the user-centered design of a decision support tool for anticoagulation in AF (ODSF step 1), built the technology to deliver this tailored decision support tool (ODSF step 2), and tested if the decision support tool with a values clarification improves the knowledge of the trade-offs of anticoagulation in AF (ODSF step 3).

## Methods

### Study Design

We used a user-centered design to develop the decision support tool. For the user-centered design, we conducted an iterative series of user experience interviews with adults recruited from the general population, medical providers, and patient-provider dyads. We recruited participants from the general Ann Arbor, Michigan, population participants during February or March 2020 (first round), April 2020 (second round), and May 2020 (third round). In addition to these general patient interviews, we interviewed 6 providers and performed 2 patient-provider dyad interviews. These patient interviews were conducted virtually due to the COVID-19 pandemic.

After completing the design of the decision support tool, we performed a randomized controlled trial using a sample of adults recruited from across the United States using a panel managed by the online survey company Qualtrics. Participants were eligible if they were 45 to 64 years old, had not been diagnosed with AF, and had not taken anticoagulants.

The Qualtrics-administered survey asked participants to imagine themselves as a patient diagnosed with AF and hypertension, which made the imaginary patient a $CHA_2DS_2$-VASc score of 1 for men and 2 for women. This was chosen because using anticoagulation in those patients is not definitive in the guidelines, and patients may need decisional support [1]. All participants then received text-based education about AF, stroke risk in AF, and the need for anticoagulation. Following the education, we randomized patients to receive no visual (standard group), a visual representation of relevant probabilities of risk of stroke in AF (visual group), or to the new decision support tool that combined design-tailored visual displays with a values clarification (visual+VC group). The survey provider performed the randomization. Quotas were used to ensure adequate sex (50% female), race (maximum of 62.3% White), and ethnicity (minimum of 12.4% not Hispanic or Latino) across all groups. Randomization was done until those quotas were met, which led to more than 200 participants in each group.

The values clarification group was presented with an exercise to evaluate which health event matters more to them: avoiding bleeding or stroke. This values clarification exercise altered the recommendation to "start anticoagulation" or "don't start anticoagulation" based on a slider movement between the 2 health events. As the user moved the slider toward avoiding a stroke, the pointer moved toward the recommendation to "start anticoagulation." As the user moved the slider toward avoiding bleeding, the pointer moved toward the recommendation to "don't start anticoagulation." In addition, those randomized to the visual or visual+VC group were subrandomized to receive either a gauge display showing the $CHA_2DS_2$-VASc score or an icon array representing the individual's probability of experiencing a stroke using a person icon [7]. The individuals' probability of experiencing a stroke did not change during the

values clarification exercise. Figures 1-4 display examples of the 4 visualizations. Participants were also asked several questions to capture baseline characteristics. The complete survey, including consent, patient scenario, educational content, and questions, is available in Multimedia Appendix 1.

**Figure 1.** Example visualization of values clarification with icon array for a 75-year-old female with hypertension.

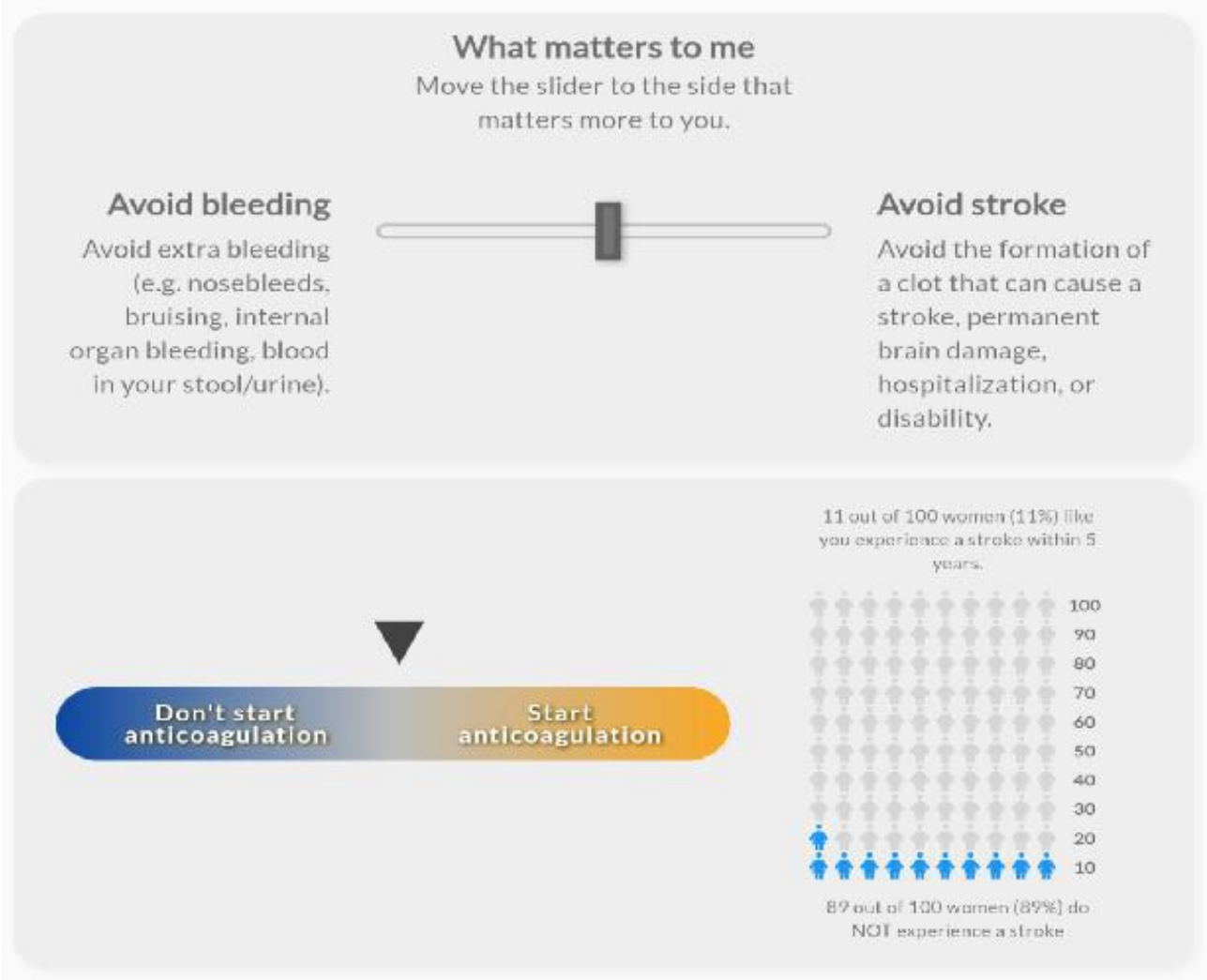


**Figure 2.** Example visualization of values clarification with gauge for a 75-year-old female with hypertension.

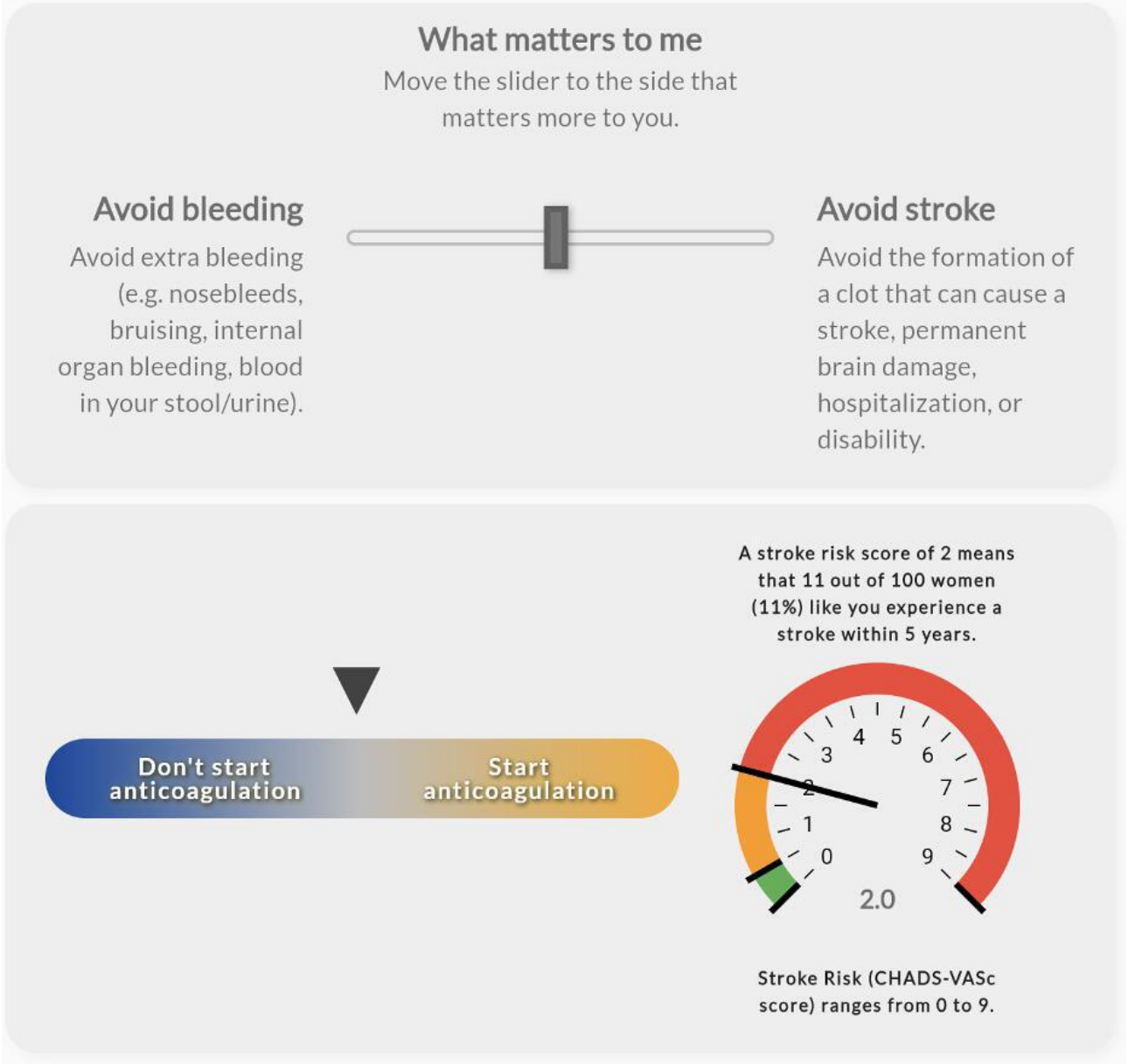


**Figure 3.** Example visualization with icon array for a 75-year-old female with hypertension.

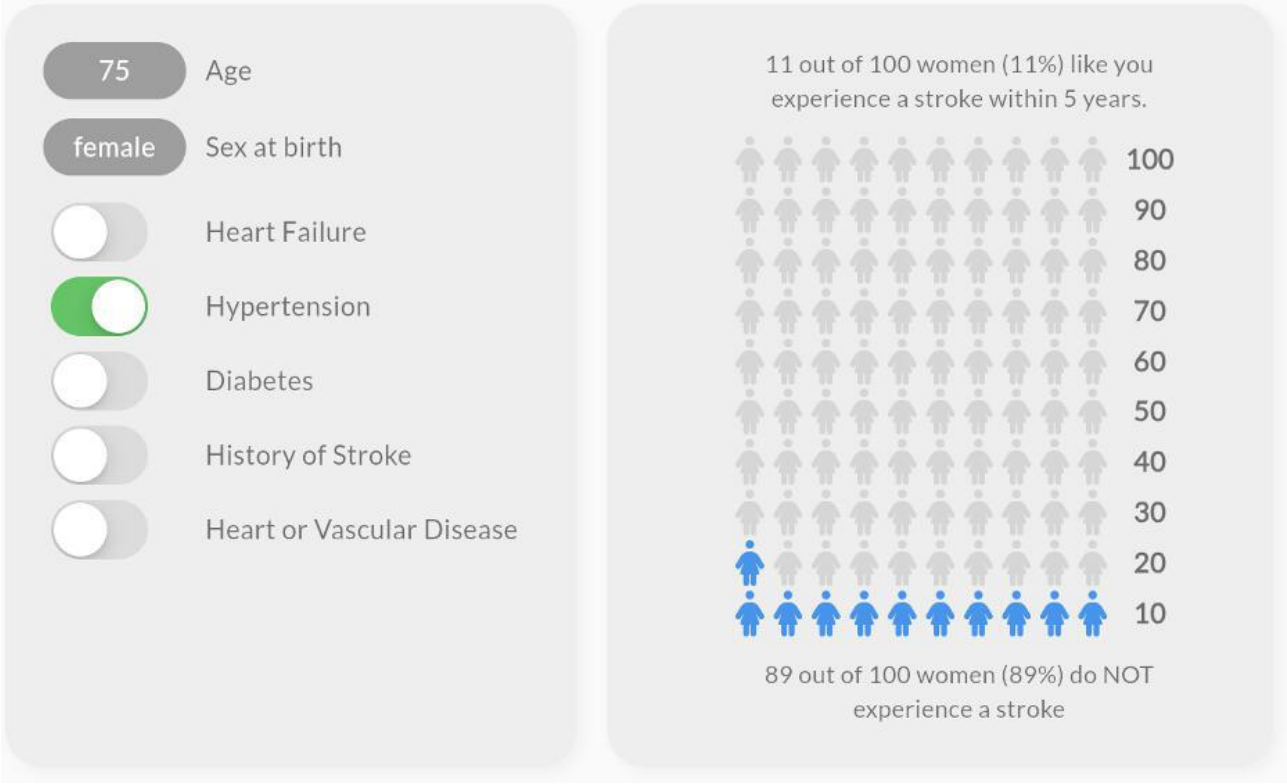



**Figure 4.** Example visualization with gauge for a 75-year-old female with hypertension.


## Outcomes

Participants completed the SURE (Sure of Myself; Understand Information; Risk-Benefit Ratio; Encouragement) screening test, which assesses the conflict a person has when making a decision [13]. The SURE test was used to understand if the participants in this study felt comfortable with their own decision to take or not take an OAC after reviewing the standard education or visuals. This was the primary outcome of this randomized trial [14]. The four yes-or-no questions are: (1) Do you feel SURE about the best choice for you? (2) Do you know the benefits and risks of each option? (3) Are you clear about

which benefits and risks matter most to you? (4) Do you have enough support and advice to make a choice? Patient comfort was assessed as the percentage of participants answering yes to all the questions. Additionally, we measured anticoagulation intentions by the question: "Based on how you feel about this decision right now, would you say you will choose to," with anchors, "Definitely TAKE an anticoagulant," (100) on the right of the scale and, "Definitely NOT take an anticoagulant," (0) on the left.

Secondary outcomes were questions about the participants' understanding of anticoagulation for SPAF. The questions were: (1) How much of a reduction would anticoagulation make to your risk of stroke in AF? (0 to 100 scale: 0=Very small to 100=Very large); (2) How important is anticoagulation for SPAF? (0 to 100 scale: 0=Not at all important to 100=Very important); (3) How worried would you be about bleeding if you took anticoagulation for SPAF? (0 to 100 scale: Not at all worried to Very worried); and (4) How worried would you be about having a stroke if you did NOT take anticoagulation? (0 to 100 scale: Not at all worried to Very worried).

### Statistical Analysis

The study was powered to detect 10 percentage differences, for example, 50% of patients in the standard group versus 60% of patients in the visual group and 70% of patients in the visual+VC group answering "Yes" to all questions on the SURE test, the primary outcome. This was considered a clinically meaningful difference between experimental groups. A total sample size of 480 survey participants (160 in each group) provided greater than 90% power to detect such a difference using a chi-square test. We set our recruitment goal for this study at 200 participants in each arm to account for variation in the estimates. The SURE test was reported as a percent of participants answering "Yes" as the numerator and the total number of participants as the denominator. The secondary outcome questions were analyzed using an analysis of variance and reported as a mean and SD of the scale in each group.

### Ethical Considerations

This study was determined to be exempt by the University of Michigan Institutional Review Board (HUM00183776). Participants consented to participate in the survey study. Completed questionnaires were collected anonymously, and the data were deidentified. The service provider, Qualtrics, was paid for each participant that completed the survey. Compensation was provided by the service provider to the participants in the study.

## Results

### Baseline Characteristics

We recruited a total of 673 participants who completed the survey and were randomized to receive standard written communication (standard group), a visual representation of relevant probabilities (visual group), or the new decision support tool that combines design-tailored visual displays with values clarification (visual+VC group). Participant enrollment and allocation are summarized in the flow diagram (Figure 5). The average age was 54 (SD 6) years, and about half of the participants in the survey were female. Table 1 shows more detailed baseline demographics of the participants.

**Figure 5.** Flow diagram for patient enrollment, randomization, and analysis.

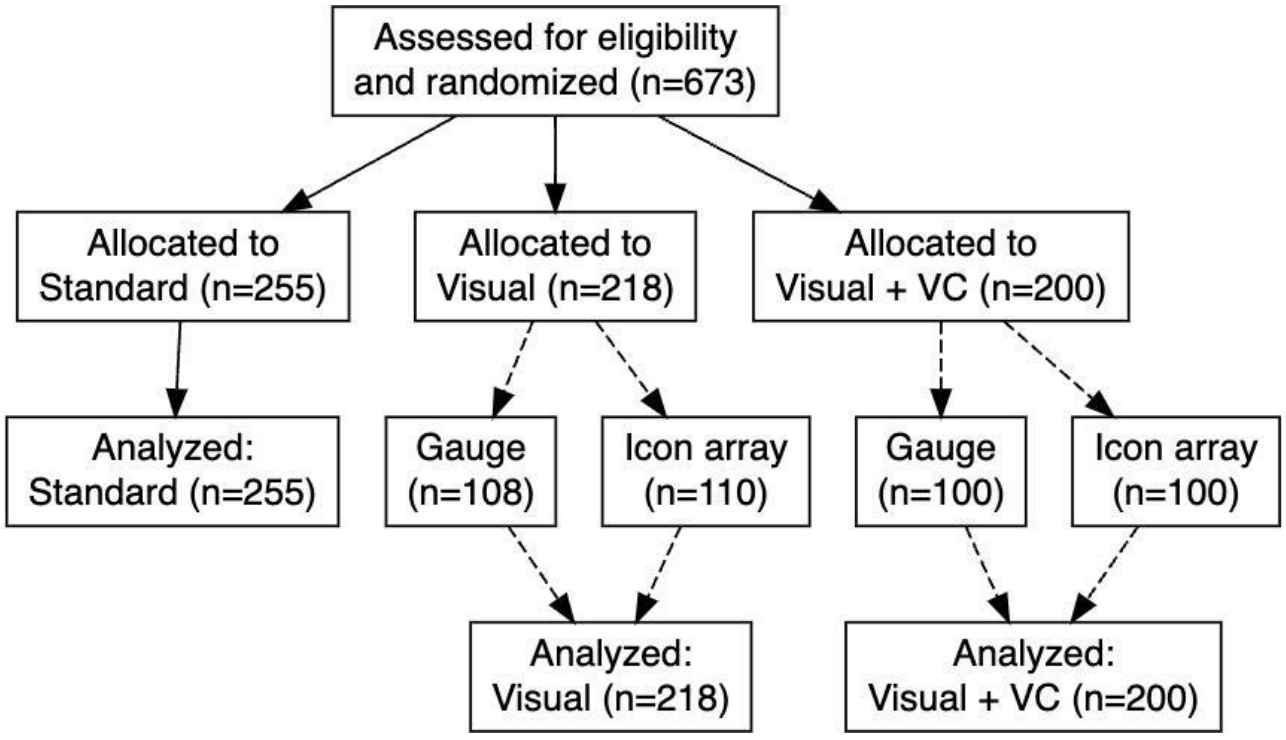


**Table .** Baseline characteristics.

| Variable | Standard (n=255) | Visual (n=218) | Visual+VC (n=200) | *P* value |
|---|---|---|---|---|
| Age (years), mean (SD) | 54.4 (5.8) | 54.5 (5.8) | 54.3 (6.1) | .93 |
| Sex (female), n (%) | 128 (50.2) | 102 (46.8) | 97 (48.5) | .76 |
| Race, n (%) | | | | .55 |
| Black | 34 (13.3) | 27 (12.4) | 26 (13) | |
| Other | 29 (11.4) | 21 (9.6) | 24 (12) | |
| White | 192 (75.3) | 170 (78) | 150 (75) | |
| Hispanic or Latino, n (%) | 55 (21.5) | 44 (20.2) | 24 (12) | .02 |
| Self-rated health status, n (%) | | | | .68 |
| Poor | 4 (1.6) | 8 (3.7) | 7 (3.5) | |
| Fair | 40 (15.7) | 43 (19.7) | 34 (17) | |
| Good | 126 (49.4) | 104 (47.7) | 90 (45) | |
| Very good | 66 (25.6) | 51 (23.4) | 57 (28.5) | |
| Excellent | 19 (7.5) | 12 (5.5) | 12 (6) | |
| Seen an HCP[a] in last 12 months, n (%) | 196 (76.9) | 162 (74.3) | 156 (78) | .66 |
| Prescription insurance, n (%) | 210 (82.4) | 177 (81.2) | 164 (82) | .95 |
| Knows someone with AFib[b], n (%) | 61 (23.9) | 64 (29.4) | 61 (30.5) | .23 |
| Knows someone taking an OAC[c], n (%) | 115 (45.1) | 103 (47.3) | 103 (51.5) | .39 |
| Confidence filling out forms, n (%) | | | | .24 |
| Never | 6 (2.4) | 3 (1.4) | 1 (0.5) | |
| Occasionally | 0 (0) | 5 (2.3) | 2 (1) | |
| Sometimes | 18 (7.1) | 11 (5.1) | 10 (5) | |
| Often | 42 (16.5) | 39 (17.9) | 40 (20) | |
| Always | 189 (74.1) | 160 (73.4) | 147 (73.5) | |
| Help reading, n (%) | 102 (40) | 74 (33.9) | 87 (43.5) | .13 |
| Problems reading, n (%) | 101 (39.6) | 77 (35.2) | 77 (38.5) | .62 |

[a]HCP: health care provider.

[b]AFib: atrial fibrillation.

[c]OAC: oral anticoagulation.

## SURE Test Results

The overall SURE test, saying “yes” to all 4 components, was 61.2% (156/255) for the standard group, 66.5% (145/218) for the visual group, and 67% (134/200) for the visual+VC group (visual vs standard, odds ratio [OR] 1.26, 95% CI 0.86 - 1.84; *P*=.23; visual+VC vs standard, OR 1.29, 95% CI 0.87 - 1.90; *P*=.20). In exploratory analyses of each question, participants felt more sure about the best choice for them, question 1 of the SURE test, if they were presented with either visual compared to standard education (visual vs standard, OR 1.59, 95% CI 1.01 - 2.49; *P*=.04; visual+VC vs standard, OR 1.48, 95% CI 0.94 - 2.33; *P*=.09). Table 2 shows the overall SURE test and the individual components.

**Table .** SURE[a] test by group.

| Variable | Standard, n (%) | Visual, n (%) | Visual+VC, n (%) | OR[b] (95% CI) and *P* value |
|---|---|---|---|---|
| Yes to all 4 SURE questions | 156 (61.2) | 145 (66.5) | 134 (67) | • Visual versus No Visual: 1.26 (0.86 - 1.84); *P*=.23<br>• Visual+VC versus No Visual: 1.29 (0.87 - 1.90); *P*=.20 |
| Do you feel SURE about the best choice for you? Yes | 191 (74.9) | 180 (82.6) | 163 (81.5) | • Visual versus No Visual: 1.59 (1.01 - 2.49); *P*=.04<br>• Visual+VC versus No Visual: 1.48 (0.94 - 2.33); *P*=.09 |
| Do you know the benefits and risks of each option? Yes | 224 (87.8) | 193 (88.5) | 179 (89.5) | • Visual versus No Visual: 1.07 (0.61 - 1.87); *P*=.82<br>• Visual+VC versus No Visual 1.18 (0.66 - 2.12); *P*=.59 |
| Are you clear about which benefits and risks matter most to you? Yes | 225 (88.2) | 185 (84.9) | 173 (86.5) | • Visual versus No Visual: 0.75 (0.44 - 1.27); *P*=.28<br>• Visual+VC versus No Visual: 0.85 (0.49 - 1.49); *P*=.58 |
| Do you have enough support and advice to make a choice? Yes | 189 (74.1) | 167 (76.6) | 151 (75.5) | • Visual versus No Visual: 1.14 (0.75 - 1.74); *P*=.53<br>• Visual + VC versus No Visual: 1.08 (0.70 - 1.65); *P*=.65 |

[a]SURE: Sure of Myself; Understand Information; Risk-Benefit Ratio; Encouragement.

[b]OR: odds ratio.

Participants were less likely to choose to take an OAC when shown either visual compared to standard education. The average rating was 58.3 (SD 30) in the standard group, 51.4 (SD 32) in the visual group, and 51.9 (SD 28) in the visual+VC group (*P*=.03). Participants also felt that the reduction in stroke risk from an OAC was less in either visual group than in the standard education group. The average rating was 63.8 (SD 22) in the standard group, 54.2 (SD 28) in the visual group, and 58.6 (SD 25) in the visual+VC group (*P*<.001). Table 3 demonstrates more detail on the questions about choosing OAC and stroke risk.

**Table .** Questions about choosing OAC[a] and stroke risk by group.

| Variable | Standard, mean (SD) | Visual, mean (SD) | Visual+VC, mean (SD) | *P* value |
|---|---|---|---|---|
| Based on how you feel about this decision right now, would you say you will choose to: 0=Do not take OAC, 100=Take OAC | 58.3 (30.0) | 51.4 (32.0) | 51.9 (28.0) | .03 |
| How much of a reduction would anticoagulation make to your risk of stroke in AFib[b]? 0=very small, 100=very large | 63.8 (22.0) | 54.2 (28.0) | 58.6 (25.0) | <.001 |
| How important is anticoagulation for stroke prevention in AFib? 0=Not important, 100=Extremely important | 75.6 (18.0) | 75.7 (19.0) | 73.9 (16.0) | .55 |
| How worried would you be about bleeding if you took anticoagulation for stroke prevention in AFib? 0=Not worried, 100=Extremely worried | 64.3 (24.0) | 65.2 (25.0) | 63 (23.0) | .63 |
| How worried would you be about having a stroke if you did NOT take anticoagulation? 0=Not worried, 100=Extremely worried | 66.3 (26.0) | 63 (28.0) | 62.1 (26.0) | .21 |

[a]OAC: oral anticoagulation.

[b]AFib: atrial fibrillation.

No significant differences were found between the visualization methods, gauge, and icon array for the outcome of the SURE test. Participants answered “yes” to all 4 SURE test questions, 65.9% (137/208) when shown a gauge and 67.6% (142/210) when shown an icon array group (*P*=.70). Participants were less likely to choose to take an OAC when shown a gauge compared to an icon array (mean 48.8, SD 31 vs mean 55.4, SD 30; *P*=.03). Participants also felt that the reduction in stroke risk from an OAC was less when shown a gauge than an icon array (mean 52.1, SD 27 vs mean 60.4, SD 25; *P*=.001). Table 4 provides further details regarding choosing OAC and stroke risk by visualization method.

**Table .** Questions about choosing OAC[a] and stroke risk by visualization method.

| Variable | Gauge (n=208), mean (SD) | Icon array (n=210), mean (SD) | *P* value |
|---|---|---|---|
| Based on how you feel about this decision right now, would you say you will choose to: 0=Do not take OAC, 100=Take OAC | 48.8 (31.0) | 55.4 (30.0) | .03 |
| How much of a reduction would anticoagulation make to your risk of stroke in AFib[b]? 0=very small, 100=very large | 52.1 (27.0) | 60.4 (25.0) | .001 |
| How important is anticoagulation for stroke prevention in AFib? 0=Not important, 100=Extremely important | 74.6 (17.0) | 75.1 (18.0) | .76 |
| How worried would you be about bleeding if you took anticoagulation for stroke prevention in AFib? 0=Not worried, 100=Extremely worried | 64.5 (24.0) | 63.7 (24.0) | .73 |
| How worried would you be about having a stroke if you did NOT take anticoagulation? 0=Not worried, 100=Extremely worried | 60.5 (27.0) | 64.7 (27.0) | .11 |

[a]OAC: oral anticoagulation.

[b]AFib: atrial fibrillation.

# Discussion

## Principal Results

This trial investigated the difference in participant preferences for OAC for SPAF after reviewing 3 different approaches, which included standard education (standard group), a visual representation of relevant probabilities of risk of stroke in AF (visual group), or the new decision support tool that combined design-tailored visual displays with a values clarification (visual+VC group). The visuals were created using a user-centered design approach with iterative feedback from patients and providers. These visuals are unique because of the addition of values clarification and because most current tools use a dot-based icon array to show stroke risk in AF [15,16]. Each participant was given a scenario with a $CHA_2DS_2$-VASc risk score, and the guidelines do not expressly state whether a patient should be prescribed an OAC. The 3 strategies did not affect the participants' comfort in deciding to take an OAC between study groups, measured by the SURE test.

Participants were less likely to take an OAC and felt that the reduction in stroke risk from an OAC was less when shown either the visual or visual VC compared to standard education. This is unique for the $CHA_2DS_2$-VASc score of 1 for men and 2 for women, which we showed participants. Since the guidelines do not recommend for or against OAC in this population, visuals like the ones in this study could persuade patients not to take OAC.

Interestingly, the values clarification visual did not demonstrate a difference in the participants' comfort in taking an OAC compared to the other visual group. This could have been due to several factors. Based on patient feedback, we used a horizontal bar for the values clarification. Previous versions of the tool we created and those in the literature used a vertical bar to represent the values clarification [8]. The horizontal bar could have led to more confusion than vertical bars. Additionally, the participants in this study were older than those in other studies using values clarification. Older participants may need more in-person help with the visuals. This could have led to more confusion with the intent of the visuals.

Although not the study's primary outcome, the 2 visual types, gauge or icon array, influenced the participants' decision to take an OAC and changed their perception of the stroke risk reduction from an OAC compared to the person-based icon array. Showing risk with the gauge made participants less likely to take an OAC, and they felt that the reduction in stroke risk from an OAC was smaller than the icon array. A body of research demonstrates the value of icon arrays in risk communication [17-20]. This difference in risk demonstration in this study could be explained by the lower detail presented in the gauge compared to the icon array, which represents a matrix of icons showing the at-risk population. The more detailed icon array could have made it easier for participants to understand the estimated risk and decide to take an OAC.

## Limitations

There are several limitations to this study. First, the tool is meant for a shared decision-making session with a patient and provider, but the survey was done with members of the general public. Second, the survey was conducted with the general public to decrease any bias the provider would add to the shared decision-making situation in the study. If this tool was implemented as shared decision-making with a provider, it could lead to a better understanding of the tool. Future research should

investigate the use of the tool with a provider present to guide and educate the patient. Third, newer AF guidelines have been published since the time of the study's completion. Although our methods and educational materials referred to earlier guidelines, the updated guidelines recognize a borderline stroke-risk threshold (eg, $CHA_2$ $DS_2$ -VASc of 1 for men or 2 for women) where shared decision-making remains a priority.

## Conclusions

Overall, the study suggests visual aids can modestly affect decision confidence and perceptions regarding the benefits of anticoagulation therapy but do not significantly change overall decision certainty in a scenario where the guidelines do not recommend for or against the treatment. Future work should determine the role of a gauge versus icon array in visual aids for decision-making in SPAF.

## Acknowledgments

The Agency for Healthcare Research and Quality funded this research (R21 HS026322).

## Data Availability

The datasets generated during or analyzed during this study are available from the corresponding author upon reasonable request.

## Authors' Contributions

MPD contributed to the conceptualization, methodology, writing – original draft, and supervision. AJF contributed to the conceptualization, methodology, and writing – review & editing. KMG and SG contributed to the writing – review & editing. GDB contributed to the conceptualization, methodology, and writing – review & editing. Finally, BZF contributed to the conceptualization, methodology, and writing - review & editing.

## Conflicts of Interest

MPD is an associate editor for *JMIR mHealth uHealth*. GDB received the following grant funding: Boston Scientific Consulting - Pfizer, Bristol-Myers Squibb, Janssen, Bayer, AstraZeneca, Sanofi, Anthos, Abbott Vascular, Boston Scientific. The authors have no further interests to declare.

Multimedia Appendix 1
Qualtrics Survey.
[PDF File, 329 KB - cardio_v9i1e67956_app1.pdf ]

## References

1. Joglar JA, Chung MK, Armbruster AL, et al. 2023 ACC/AHA/ACCP/HRS guideline for the diagnosis and management of atrial fibrillation: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. Circulation 2024 Jan 2;149(1):e1-e156. [doi: 10.1161/CIR.0000000000001193] [Medline: 38033089]
2. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007 Jun 19;146(12):857-867. [doi: 10.7326/0003-4819-146-12-200706190-00007] [Medline: 17577005]
3. Noseworthy PA, Brito JP, Kunneman M, et al. Shared decision-making in atrial fibrillation: navigating complex issues in partnership with the patient. J Interv Card Electrophysiol 2019 Nov;56(2):159-163. [doi: 10.1007/s10840-018-0465-5] [Medline: 30327992]
4. Lip GYH, Frison L, Halperin JL, Lane DA. Identifying patients at high risk for stroke despite anticoagulation: a comparison of contemporary stroke risk stratification schemes in an anticoagulated atrial fibrillation cohort. Stroke 2010 Dec;41(12):2731-2738. [doi: 10.1161/STROKEAHA.110.590257] [Medline: 20966417]
5. van den Ham HA, Klungel OH, Singer DE, Leufkens HGM, van Staa TP. Comparative performance of ATRIA, CHADS2, and CHA2DS2-VASc risk scores predicting stroke in patients with atrial fibrillation: results from a National Primary Care Database. J Am Coll Cardiol 2015 Oct 27;66(17):1851-1859. [doi: 10.1016/j.jacc.2015.08.033] [Medline: 26493655]
6. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns H, Lip GYH. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. Chest 2010 Nov;138(5):1093-1100. [doi: 10.1378/chest.10-0134] [Medline: 20299623]
7. Zikmund-Fisher BJ, Witteman HO, Dickson M, et al. Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. Med Decis Making 2014 May;34(4):443-453. [doi: 10.1177/0272989X13511706] [Medline: 24246564]
8. Witteman HO, Chipenda Dansokho S, Exe N, Dupuis A, Provencher T, Zikmund-Fisher BJ. Risk communication, values clarification, and vaccination decisions. Risk Anal 2015 Oct;35(10):1801-1819. [doi: 10.1111/risa.12418] [Medline: 25996456]

9. Witteman HO, Scherer LD, Gavaruzzi T, et al. Design features of explicit values clarification methods: a systematic review. Med Decis Making 2016 May;36(4):453-471. [doi: 10.1177/0272989X15626397] [Medline: 26826032]
10. Witteman HO, Gavaruzzi T, Scherer LD, et al. Effects of design features of explicit values clarification methods: a systematic review. Med Decis Making 2016 Aug;36(6):760-776. [doi: 10.1177/0272989X16634085] [Medline: 27044883]
11. Murray MA, Miller T, Fiset V, O'Connor A, Jacobsen MJ. Decision support: helping patients and families to find a balance at the end of life. Int J Palliat Nurs 2004 Jun;10(6):270-277. [doi: 10.12968/ijpn.2004.10.6.13268] [Medline: 15284621]
12. O'Connor AM, Tugwell P, Wells GA, et al. A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. Patient Educ Couns 1998 Mar;33(3):267-279. [doi: 10.1016/s0738-3991(98)00026-3] [Medline: 9731164]
13. Légaré F, Kearing S, Clay K, et al. Are you SURE?: Assessing patient decisional conflict with a 4-item screening test. Can Fam Physician 2010 Aug;56(8):e308-e314. [Medline: 20705870]
14. Parayre AF, Labrecque M, Rousseau M, Turcotte S, Légaré F. Validation of SURE, a four-item clinical checklist for detecting decisional conflict in patients. Med Decis Making 2014 Jan;34(1):54-62. [doi: 10.1177/0272989X13491463] [Medline: 23776141]
15. Kunneman M, Branda ME, Hargraves IG, et al. Assessment of shared decision-making for stroke prevention in patients with atrial fibrillation: a randomized clinical trial. JAMA Intern Med 2020 Sep 1;180(9):1215-1224. [doi: 10.1001/jamainternmed.2020.2908] [Medline: 32897386]
16. Noseworthy PA, Branda ME, Kunneman M, et al. Effect of shared decision-making for stroke prevention on treatment adherence and safety outcomes in patients with atrial fibrillation: a randomized clinical trial. J Am Heart Assoc 2022 Jan 18;11(2):e023048. [doi: 10.1161/JAHA.121.023048] [Medline: 35023356]
17. Hawley ST, Zikmund-Fisher B, Ubel P, Jancovic A, Lucas T, Fagerlin A. The impact of the format of graphical presentation on health-related knowledge and treatment choices. Patient Educ Couns 2008 Dec;73(3):448-455. [doi: 10.1016/j.pec.2008.07.023] [Medline: 18755566]
18. Tait AR, Voepel-Lewis T, Zikmund-Fisher BJ, Fagerlin A. The effect of format on parents' understanding of the risks and benefits of clinical research: a comparison between text, tables, and graphics. J Health Commun 2010 Jul;15(5):487-501. [doi: 10.1080/10810730.2010.492560] [Medline: 20677054]
19. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. Health Psychol 2009 Mar;28(2):210-216. [doi: 10.1037/a0014474] [Medline: 19290713]
20. Garcia-Retamero R, Galesic M. Who proficts from visual aids: overcoming challenges in people's understanding of risks. Soc Sci Med 2010 Apr;70(7):1019-1025. [doi: 10.1016/j.socscimed.2009.11.031]


## Abbreviations

**AF:** atrial fibrillation
**$CHA_2DS_2$-VASc:** congestive heart failure; hypertension; age ≥75 years [doubled]; type 2 diabetes; previous stroke, transient ischemic attack, or thromboembolism [doubled]; vascular disease; age 65 to 75 years; and sex category
**HAS-BLED:** hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile international normalized ratio, elderly [>65 years], drugs/alcohol concomitantly
**OAC:** oral anticoagulation
**ODSF:** Ottawa Decision Support Framework
**OR:** odds ratio
**SPAF:** stroke prevention in atrial fibrillation
**SURE:** Sure of Myself; Understand Information; Risk-Benefit Ratio; Encouragement
**visual+VC:** visual aids plus a values clarification exercise


*Edited by A Coristine; submitted 24.10.24; peer-reviewed by A Allen, E Kodani; revised version received 18.02.25; accepted 18.02.25; published 11.04.25.*

*Please cite as:*
*Dorsch MP, Flynn AJ, Greer KM, Ganai S, Barnes GD, Zikmund-Fisher B*
*A Web-Based Tool to Perform a Values Clarification for Stroke Prevention in Patients With Atrial Fibrillation: Design and Preliminary Testing Study*
*JMIR Cardio 2025;9:e67956*
*URL: https://cardio.jmir.org/2025/1/e67956*
*doi:10.2196/67956*

Original Paper

# Acceptability of a Web-Based Health App (PortfolioDiet.app) to Translate a Nutrition Therapy for Cardiovascular Disease in High-Risk Adults: Mixed Methods Randomized Ancillary Pilot Study

Meaghan E Kavanagh[1,2], MSc, PhD; Laura Chiavaroli[1,2,3], MSc, PhD; Selina M Quibrantar[1], MSc; Gabrielle Viscardi[1], BSc (Hons), RD; Kimberly Ramboanga[2], BASc; Natalie Amlin[2], BSc, RD; Melanie Paquette[1,2], RD, MSc; Sandhya Sahye-Pudaruth[1,2], RD, MSc; Darshna Patel[1,2], BA; Shannan M Grant[4,5], PDt, RD, MSc, PhD; Andrea J Glenn[6,7], RD, PhD; Sabrina Ayoub-Charette[1,2], BSc; Andreea Zurbau[1,2], RD, PhD; Robert G Josse[8], MD; Vasanti S Malik[1,7], PhD; Cyril W C Kendall[1,2], PhD; David J A Jenkins[1,2,3,8,9], MD, PhD; John L Sievenpiper[1,2,3,8,9], MD, PhD

[1]Department of Nutritional Sciences, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[2]Toronto 3D Knowledge Synthesis and Clinical Trials Unit, Clinical Nutrition and Risk Factor Modification Center, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

[3]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

[4]Departments of Pediatrics and Obstetrics and Gynaecology, IWK Health, Halifax, NS, Canada

[5]Department of Obstetrics and Gynaecology, Faulty of Medicine, Dalhousie University, Halifax, NS, Canada

[6]Department of Nutrition and Food Studies, New York University, New York, NY, United States

[7]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, United States

[8]Division of Endocrinology and Metabolism, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

[9]Department of Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

**Corresponding Author:**
John L Sievenpiper, MD, PhD
Department of Nutritional Sciences
Temerty Faculty of Medicine
University of Toronto
C. David Naylor Building, Room 201
6 Queen's Park Crescent
Toronto, ON, M5S 3H2
Canada
Phone: 1 416 867 3732
Fax: 1 416 867 7495
Email: john.sievenpiper@utoronto.ca

## Abstract

**Background:** The Portfolio Diet is a dietary pattern for cardiovascular disease (CVD) risk reduction with 5 key categories including nuts and seeds; plant protein from specific food sources; viscous fiber sources; plant sterols; and plant-derived monounsaturated fatty acid sources. To enhance implementation of the Portfolio Diet, we developed the PortfolioDiet.app, an automated, web-based, multicomponent, patient-facing health app that was developed with psychological theory.

**Objective:** We aimed to evaluate the effect of the PortfolioDiet.app on dietary adherence and its acceptability among adults with a high risk of CVD over 12 weeks.

**Methods:** Potential participants with evidence of atherosclerosis and a minimum of one additional CVD risk factor in an ongoing trial were invited to participate in a remote web-based ancillary study by email. Eligible participants were randomized in a 1:1 ratio using a concealed computer-generated allocation sequence to the PortfolioDiet.app group or a control group for 12 weeks. Adherence to the Portfolio Diet was assessed by weighed 7-day diet records at baseline and 12 weeks using the clinical Portfolio Diet Score, ranging from 0 to 25. Acceptability of the app was evaluated using a multifaceted approach, including usability

through the System Usability Scale ranging from 0 to 100, with a score >70 being considered acceptable, and a qualitative analysis of open-ended questions using NVivo 12.

**Results:** In total, 41 participants were invited from the main trial to join the ancillary study by email, of which 15 agreed, and 14 were randomized (8 in the intervention group and 6 in the control group) and completed the ancillary study. At baseline, adherence to the Portfolio Diet was high in both groups with a mean clinical Portfolio Diet Score of 13.2 (SD 3.7; 13.2/25, 53%) and 13.7 (SD 5.8; 13.7/25, 55%) in the app and control groups, respectively. After the 12 weeks, there was a tendency for a mean increase in adherence to the Portfolio Diet by 1.25 (SD 2.8; 1.25/25, 5%) and 0.19 (SD 4.4; 0.19/25, 0.8%) points in the app and control group, respectively, with no difference between groups ($P$=.62). Participants used the app on average for 18 (SD 14) days per month and rated the app as usable (System Usability Scale of mean 80.9, SD 17.3). Qualitative analyses identified 4 main themes (user engagement, usability, external factors, and added components), which complemented the quantitative data obtained.

**Conclusions:** Although adherence was higher for the PortfolioDiet.app group, no difference in adherence was found between the groups in this small ancillary study. However, this study demonstrates that the PortfolioDiet.app is considered usable by high-risk adults and may reinforce dietitian advice to follow the Portfolio Diet when it is a part of a trial for CVD management.

**Trial Registration:** ClinicalTrials.gov NCT02481466; https://clinicaltrials.gov/study/NCT02481466

***(JMIR Cardio 2025;9:e58124)*** doi:10.2196/58124

**KEYWORDS**

diet; apps; dietary app; Portfolio Diet; dietary portfolio; cholesterol reduction; cardiovascular disease; eHealth; usability; acceptability

## Introduction

### Background

Cardiovascular disease (CVD) remains the leading cause of death globally [1]. Effective prevention and management strategies are needed to target modifiable risk factors for CVD. Several recent Canadian population-based studies have shown that many patients at high CVD risk continue to have low-density lipoprotein cholesterol (LDL-C) levels well above the guideline targets [2,3]. LDL-C has been extensively studied and described as a causal factor for CVD [4]. LDL-C levels above the target can result from multiple factors such as insufficient LDL-C lowering with statins, statin-related side effects, suboptimal medication adherence, and treatment inertia [5]. Amid these challenges, dietary approaches for CVD risk reduction emerge as a potentially powerful tool [6] with clinical practice guidelines universally recommending diet and lifestyle as the cornerstone of therapy for addressing CVD [7,8].

The Portfolio Diet is a dietary pattern recognized by clinical practice guidelines in Canada and internationally, including the Canadian Cardiovascular Society [7,8] Diabetes Canada [9], Obesity Canada [10], Canadian Cardiovascular Harmonized National Guidelines Endeavour [11], Heart UK [12], European Atherosclerosis Society [13], and the American College of Cardiology and American Heart Association guidelines [14]. The Portfolio Diet has been shown to have the same LDL-C and inflammatory (C-reactive protein) reductions (approximately 30%) as statin therapy in a head-to-head randomized controlled trial in participants with hyperlipidemia [15]. A systematic review and meta-analysis of clinical trials [16] confirmed these “drug-like” effects and demonstrated clinically meaningful cardiovascular benefits on other targets including non–high-density lipoprotein cholesterol, apolipoprotein B, triglycerides, blood pressure, and estimated 10-year CVD risk.

Although the Portfolio Diet, among other dietary patterns, is recognized in guidelines, uptake and implementation of nutrition therapies in clinical practice remains limited. This dilemma stems from several barriers that hinder the widespread adoption of nutrition therapies. Chief among these challenges are the shortage of available health support services, the restricted access to registered dietitians, the time constraints faced by physicians, and the lack of comprehensive education and tools [17,18]. The resulting consequence of these obstacles is that many patients who would benefit from nutrition therapy do not receive it [19]. In a survey of Canadian patients randomly selected from family health networks, only 37% reported receiving nutrition counseling in primary care [20], highlighting the need for effective dissemination strategies.

Due to their highly scalable nature, the use of technology to aid in the dissemination and delivery of lifestyle behavior change interventions has become of great interest with the number of studies investigating health apps having gone up rapidly since 2010 [21]. Web- and mobile-based applications (hereafter apps) provide an important alternative and complementary approach for the delivery and long-term reinforcement of health advice. Previous work has found that apps can be a cost-effective method for the delivery of lifestyle interventions such as in smoking cessation [22,23]. As smartphones become common everyday household items, the possible reach and impact of using apps to deliver interventions grows. Currently, it is estimated that over 300,000 health apps exist on app stores [24]; however, most publicly available health apps remain untested.

### Objective

To enhance the implementation of the Portfolio Diet in health care settings, we developed the PortfolioDiet.app [25], a free, web-based, multicomponent, patient-facing engagement and educational tool. While we have previously undertaken quality improvement and usability testing of the PortfolioDiet.app in a convenience sample [26], the app has not yet been evaluated in its intended population of adults at high risk of CVD. Therefore, the objective of this study was to evaluate the effect of the PortfolioDiet.app on dietary adherence and its

acceptability among adults with a high risk of CVD over 12 weeks.

## Methods

### Design

This mixed methods ancillary study was a 12-week single-center, open label, randomized controlled ancillary study within an ongoing 3-year multicenter randomized controlled trial, the Combined Portfolio Diet and Exercise Study (PortfolioEx; ClinicalTrials.gov NCT02481466). All participants for the ancillary study were recruited from those randomized to one of the 2 Portfolio Diet arms at the St. Michael's Hospital, Toronto, Canada, site of the main trial. Recruited participants were randomized to receive the PortfolioDiet.app for 12 weeks or to the control group.

We used a mixed methods approach where we collected and analyzed both quantitative and qualitative data to ensure a thorough and comprehensive assessment of the intervention [27]. Multimedia Appendix 1 shows the CONSORT (Consolidated Standards of Reporting Trials) checklist and Multimedia Appendix 2 shows the CONSORT - EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (version 1.6.1) [28]. Our intention was to gather complementary data from quantitative and qualitative methods and then integrate findings within a data triangulation design [29], enabling us to draw metainferences regarding the acceptability and usability of the PortfolioDiet.app. These insights will not only inform potential refinements to the app itself but also guide the design of a future trial.

### Ethical Considerations

The ancillary study was conducted according to the guidelines of the Declaration of Helsinki and approved as an ancillary study to the main PortfolioEx trial by the research ethics board (REB) of St. Michael's Hospital, Unity Health Toronto (REB 14-316). All participants provided written informed consent to the main trial and separately provided verbal over-the-phone informed consent to the ancillary study. No compensation was provided to the participants. Participant data were stored at St. Michael's Hospital and kept confidential by ensuring identifying data were kept separate from study data using a study ID. All study data were deidentified, and the master linking log was kept in a password-protected file on a secure drive at St. Michael's Hospital, only accessible to study staff.

### Study Participants

Participants in the PortfolioEx trial were men and postmenopausal women with a BMI ≤40 kg/m$^2$ who were considered at high risk for CVD. Participants had carotid artery plaque buildup (an intima-media thickness of ≥1.2 mm) in addition to at least one other of the following characteristics: type 2 diabetes, history of myocardial infarction or angioplasty, hypercholesterolemia and treated with statins or prescribed statins but due to statin intolerance or choice are not taking statins, or raised blood pressure (>140/90 mm Hg). To be eligible for the ancillary study, participants needed to have access to a web portal through a personal smartphone, tablet, or home computer and needed to have an active email address.

### Randomization

Eligible participants were randomized in a 1:1 ratio to either the PortfolioDiet.app group or a control group using a computer-generated allocation sequence. Randomization was done using block sizes of 4 with stratification by sex (ie, male and female), age (ie, <65 and ≥65 years), and their allocated exercise group (ie, yes and no) in the 3-year PortfolioEx trial. MK was responsible for contacting and enrolling participants, providing them with information on the study, and sending them app details and questionnaires. The randomization table was developed from Sealed Envelope [30]. SA-C, who had no contact with participants, was the only one to have access to the randomization table and was responsible for assigning participants to the interventions.

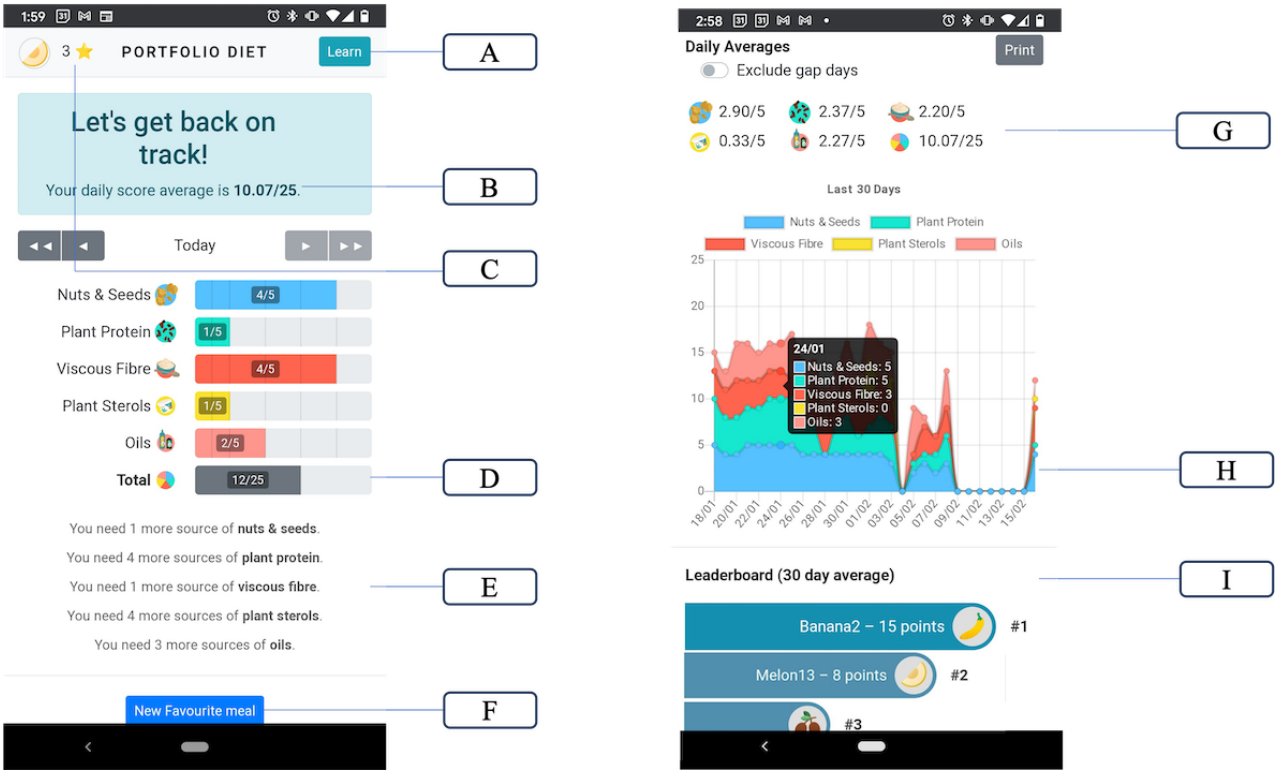
### Theoretical and Operational Design Components of the PortfolioDiet.app

The app was developed with integration of the psychological theory including the social cognitive theory and self-regulatory principles of behavior change by providing multiple forms of behavioral feedback on dietary adherence, including tip sheets for promoting dietary change. Designed with a variety of elements to enhance and sustain behavior change, Figure 1 shows the PortfolioDiet.app's home page with key features highlighted. These include features previously identified as elements preferred by health app users, including a personalized dashboard, goal setting, educational features, and email messages.

Within the "Learn" section, the app houses educational resources including a bank of recipes, tip sheets, and videos (Figure 1). The PortfolioDiet.app offers users an array of recipes that span from family friendly dinner recipes to quick snacks while also including culturally adapted recipes (eg, African, Mediterranean, and South Asian) and filters for dietary restrictions (eg, gluten free, low carbohydrates, and low sodium).

Many of the features would fall under the definition of gamification, which evidence from a systematic review and meta-analysis has found to support behavior change, increasing measures of physical activity and decreasing measures of adiposity [31]. These features include star rewards for engaging with the app, allowing users to track and visualize their average adherence and progress, provides daily goals, and a social competitive aspect through a leaderboard on diet adherence (Figure 1). Star rewards allow users to earn and collect stars, incentivizing users to interact with the app. Users can collect stars by logging into the app and correctly answering the question of the week. The leaderboard feature provides users with an overview of the number of app members and their average daily score over 30 days, allowing users to compare their average daily score with other users.

**Figure 1.** PortfolioDiet.app dashboard with key features highlighted, top to bottom: (A) learn tab that has recipes, tipsheets, and videos; (B) daily average Portfolio Diet Score per month; (C) star rewards, a form of reward for logging into the app and for completing the question of the week; (D) current day total Portfolio Diet Score; (E) specific daily messages related to goals; (F) personal favorite meals for easy tracking; (G) subcategory Portfolio Diet Scores with daily targets of 5 points; (H) progress-tracking graph displaying the monthly progress of the score; and (I) leaderboard with other app users' 30-day average.




The app uses a dietary score to guide users on the amount of key foods to eat and to provide personalized feedback. The clinical Portfolio Diet Score (c-PDS) has previously been validated in a similar population of adults with hyperlipidemia [32]. By following the Portfolio Diet, users can earn up to 5 points from each category of Portfolio Diet foods for a maximum c-PDS of 25 points per day in the app. It has previously been shown that an increase in c-PDS by 12 points predicts about a 0.53 mmol/L (13%) reduction in LDL-C in patients with hyperlipidemia over 6 months [32].

When using the PortfolioDiet.app, users can input Portfolio Diet foods and portion sizes. Each food item is shown as 1 portion size, in grams or as cup measurements, with targets based on 1 of 3 caloric levels. The user picks the portion size that is most similar to their intake and then the item will appear on their dashboard. The app allows for self-monitoring and provides feedback through an average daily score on the home page and the current day's score and, below, a graph displaying the monthly progress for dietary adherence (Figure 1).

## Intervention

Participants randomized to receive the PortfolioDiet.app were sent an instructional guide and videos by email, with instructions on how to create an account and use the app features. The PortfolioDiet.app is fully automated and was provided as a web-based version that could be used on laptops, tablets, smartphones, or public computers such as those in libraries. The dietary advice on the Portfolio Diet conveyed through the app included recommendations on the 5 core cholesterol-lowering foods and their recommended amounts per day for a 2000 kcal diet: 45 g nuts and seeds (eg, tree nuts, peanuts, or seeds); 50 g of plant protein (eg, from soy and dietary pulses); 20 g viscous soluble fiber (eg, from sources such as oats, barley, psyllium, eggplant, okra, apples, oranges, or berries); 2 g plant sterols (eg, from supplements and fortified foods); and 45 g monounsaturated fatty acids (eg, from cold-pressed olive, canola, soybean, "high-oleic" sunflower and safflower oils, or avocados).

Development of the app was frozen during the trial. Participants randomized to the PortfolioDiet.app group were asked to use the app every day (ie, including both weekdays and weekends) over the 12-week intervention in the ancillary study. If a day was missed, participants were encouraged in the app to retroactively enter their Portfolio Diet foods. If participants did not make an account during the first week, they were sent an email reminder every week. Participants were not blinded to their allocation and neither were the study staff. Participants randomized to the control group were informed of their randomization allocation but received no further contact from the PortfolioDiet.app staff until after the study, at which point they were offered access to the app. The 12-week intervention length was chosen to allow for a controlled assessment of the health app on dietary adherence (the main objective), without unfairly restricting access to the app for those participants randomized to not receive the app within an active trial.

As REB approval for this ancillary study was received during the COVID-19 publicly declared emergency (ie, the pandemic). Staff were not permitted to access Unity Health sites or to have in-person contact with participants or staff. Therefore, all study

interactions with participants for the study took place over the phone or by email. The interactions in the ancillary study did not provide any dietary counseling support and only provided minimal-contact administrative support to those randomized to the PortfolioDiet.app group, including encouraging the use of materials provided to help with account creation and using the app features.

### Outcomes

The primary outcome was a change in dietary adherence to the Portfolio Diet over 12 weeks in those randomized to the PortfolioDiet.app group compared to those in the control group. Adherence to the Portfolio Diet was assessed from weighed 7-day diet records (7DDRs) collected at baseline and at 12 weeks through predesigned paper-based templates. Participants were trained and supported by registered dietitians to complete the records, and paper copies were mailed to participants with telephone discussions scheduled every 3 months. The c-PDS was calculated from the 7DDRs and ranges from 0 to 25 points, with a score of 0 indicating no adherence to the Portfolio Diet and a score of 25 indicating full adherence to the diet.

Acceptability of the PortfolioDiet.app was assessed in participants who were randomized to the PortfolioDiet.app group. App use was evaluated through the app's web-based repository based on participants' log-ins over the 12 weeks. Usability was assessed using the System Usability Scale (SUS). The SUS is a validated usability questionnaire that has been used in clinical settings to assess the usability of various systems and tools [33,34]. The SUS includes 10 statements rated on a 5-point Likert scale (from 1=strongly disagree to 5=strongly agree). The score ranges from 0 to 100 with a score higher than 70 being considered acceptable [35]. We also collected the c-PDS data from the app, which were based on participants' logged entries into the app. The c-PDS was saved in the app's web-based repository and, unlike the primary outcome of dietary adherence, was not calculated from the 7DDR.

Multimedia Appendix 3 shows the structured questionnaire used with open-ended questions. The questionnaire collected participant feedback on acceptability, knowledge acquisition, and app features. It was developed by MEK, LC, and SMG using existing tools [36] and included the SUS questionnaire [33]. The questionnaire was emailed to participants after 12 weeks of using the PortfolioDiet.app. Participants were instructed to complete the questionnaire by typing out their responses and to return it by email.

### Analytic Techniques

As part of the primary 3-year PortfolioEx trial, eligibility by intima-media thickness was measured by B-mode Carotid Ultrasound at 12 carotid artery segments (1-cm long) of the near and far walls of the internal, bifurcation, and common left and right carotid arteries. Baseline serum lipids were measured on fasting serum and analyzed in the routine hospital laboratory using Beckman SYNCHRON LX Systems. The LDL-C level was calculated using the Friedewald equation [37]. Anthropometric data were collected when participants were fasting by trained study staff, and information on medications and the diagnosis of type 2 diabetes was collected through self-report questionnaires.

### Analyses

Baseline characteristics were assessed by 2-sample *t* tests for continuous variables and Fisher exact test for categorical variables. Dietary adherence to the Portfolio Diet from weighed 7DDRs measured by the c-PDS (week 0 to week 12) was expressed as mean differences with SDs. Within-group and between-group differences were assessed using a 2-sample *t* test. On the basis of a prior multi-center randomized controlled trial, a total of 56 participants were required to detect a ≥3.28 point difference in c-PDS with 80% power (1-β), =.05, and SD 4.30 [38]. Statistical analysis was performed using Stata (version 7; StataCorp). The planned sample size of 56 participants, with approximately 23 receiving the app, was deemed sufficient to reach data saturation, particularly given our homogeneous study population, and aligned with the study by Hennink and Kaiser [39], who suggest that smaller sample sizes can be adequate for achieving saturation in qualitative research with homogeneous groups.

For the qualitative analysis, open-ended survey data were extracted from the structured questionnaire and initially analyzed independently using NVivo (version 12.7.0; QSR International) by members of the research team (MEK, LC, SMQ, and GV). The team used reflexive thematic analysis, as described in the study by Braun and Clarke [40], to identify patterns and concepts within the data [40]. A coding framework was collaboratively established, and each member of the research team conducted an individual review of both the data and the coding framework to confirm the accuracy of the interpretations during initial analysis, and to identify any elements or insights that might not have been initially captured during the group analysis. Regular team meetings were held weekly over a 1-month period to discuss coding findings, address discrepancies, and reach consensus on the identified codes. Identified codes were further structured into main themes and subthemes, and a table was produced to arrange quotations derived from the survey responses to substantiate the themes and subthemes identified.

The analysis process was performed with consideration of the trustworthiness criteria [40]. To ensure credibility, dependability, confirmability, and transferability in the qualitative analysis, multiple researchers were involved in the coding process to reach consensus on identified themes, a detailed description of coding decisions and theme development was maintained, and potential biases were acknowledged with regular discussions to minimize influence. In addition, a detailed description of the study, participants, and findings was provided to enable readers to assess the applicability of the results to other settings.

After both analyses were conducted, the qualitative findings were compared with the quantitative results using a data triangulation approach.

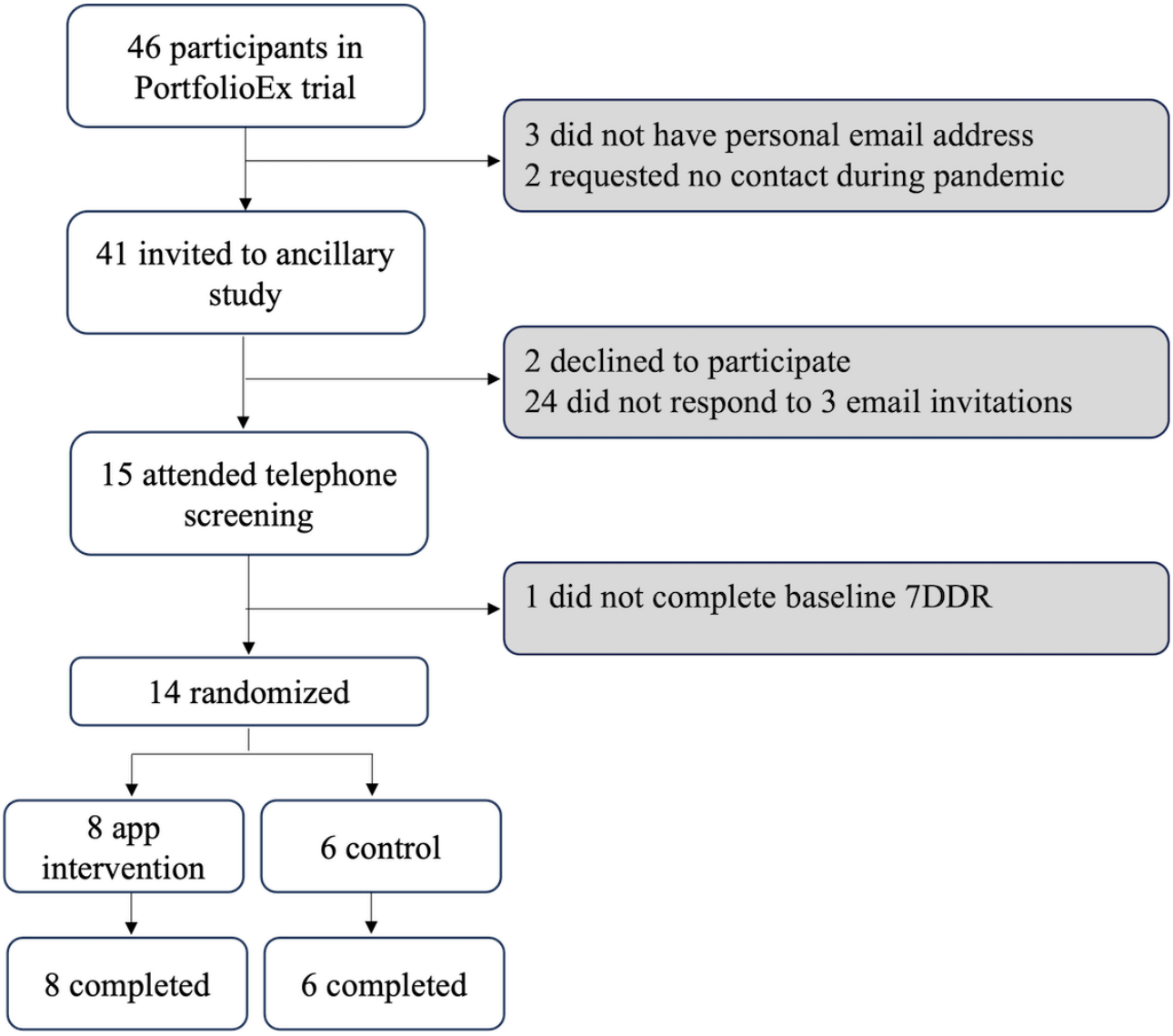
## Results

### Consolidated Standards of Reporting Trials

Figure 2 shows the CONSORT diagram of participants in the ancillary study. While a total of 66 participants were randomized

to the PortfolioDiet.app group arm in the PortfolioEx trial, 14 dropped out before the ancillary study began. Once REB approval was received, 6 participants had completed the trial or were scheduled to complete the trial within 3 months. Therefore, of the remaining 46 participants, 41 were eligible (3 did not have a personal email and 2 requested no contact during the COVID-19 pandemic). Potential participants were invited by email to participate in the ancillary study. Between July 2021 and February 2022, of the 15 participants who completed a telephone screen, 14 had baseline dietary data and were randomized (intervention group: n=8; control group: n=6) and completed the study. The average duration that the 14 participants had been enrolled in the PortfolioEx trial and were receiving the Portfolio Diet intervention at the Toronto site was 24.6 (SD 4.1; range 18-33) months.

**Figure 2.** CONSORT (Consolidated Standards of Reporting Trials) flow diagram showing participant flow through the ancillary study. PortfolioEx trial: the Combined Portfolio Diet and Exercise Study; 7DDR: 7-day diet record.




## Baseline Characteristics

Table 1 shows the baseline characteristics of the 14 randomized participants. Participants were primarily female (n=9, 64%), identified mostly as White (n=7, 50%) followed by South Asian (n=3, 21%), Filipino (n=2, 14%), and Black (n=2, 14%). Their mean age was 65 (SD 4, range 52-79) years; 71% (10/14) were on lipid-lowering medication and 29% (4/14) had type 2 diabetes. Adherence to the Portfolio Diet was high in both groups at baseline with a c-PDS of 53% (13.2/25) in the app group and 55% (13.7/25) in the control group. A total of 2 participants (1 in the app group and 1 in the control group) did not provide their final 12-week 7DDR. Therefore, they were excluded from the primary analysis.

**Table 1.** Baseline characteristics of participants.

| | Total (N=14) | App group (n=8) | Control group (n=6) | *P* value |
|---|---|---|---|---|
| Age (y), mean (SD) | 65.4 (9) | 65 (9) | 66 (9) | .96 |
| **Sex, n (%)** | | | | .30 |
| Female | 9 (64) | 4 (50) | 5 (83) | |
| Male | 5 (36) | 4 (50) | 1 (17) | |
| **Race or ethnicity, n (%)** | | | | .99 |
| Black | 2 (14) | 1 (12.5) | 1 (17) | |
| Filipino | 2 (14) | 1 (12.5) | 1 (17) | |
| South Asian | 3 (21) | 2 (25) | 1 (17) | |
| White | 7 (50) | 4 (50) | 3 (50) | |
| Body weight (kg), mean (SD) | 73.1 (13.3) | 68.9 (11.3) | 78.7 (14.8) | .18 |
| BMI (kg/m$^2$), mean (SD) | 28.0 (4.3) | 26.2 (4.5) | 30.3 (5.1) | .07 |
| Waist circumference (cm), mean (SD) | 96.6 (11.5) | 92.1 (10.7) | 102.7 (10.3) | .09 |
| **BP[a] (mm Hg), mean (SD)** | | | | |
| Systolic BP | 114.8 (16.9) | 113.8 (11.7) | 116.1 (23.4) | .80 |
| Diastolic BP | 64.9 (11.4) | 61 (7.3) | 70 (14.5) | .15 |
| Type 2 diabetes, n (%) | 4 (29) | 2 (25) | 2 (33) | .99 |
| **Lipids (mmol/L), mean (SD)** | | | | |
| Total cholesterol | 4.8 (1.7) | 4.6 (1.9) | 5.0 (1.4) | .75 |
| LDL-C[b,c] | 2.8 (1.5) | 2.7 (1.9) | 2.9 (1.1) | .79 |
| HDL-C[d] | 1.4 (0.4) | 1.5 (0.3) | 1.4 (0.4) | .70 |
| Non-HDL-C | 3.4 (1.6) | 3.3 (2.0) | 3.5 (1.2) | .83 |
| Triglycerides | 1.3 (0.5) | 1.3 (0.6) | 1.3 (0.4) | .99 |
| **Medication use** | | | | |
| Lipid-lowering medication, n (%) | 10 (71) | 7 (88) | 3 (50) | .25 |
| Antihypertensive medication, n (%) | 9 (64) | 5 (63) | 4 (67) | .99 |
| Duration enrolled in the PortfolioEx trial (months), mean (SD) | 24.6 (4.1) | 23.3 (4.7) | 26.5 (2.3) | .15 |
| c-PDS[e] (points; range 0 to 25), mean (SD) | 13.4 (4.4) | 13.2 (3.7) | 13.7 (5.8) | .87 |

[a]BP: blood pressure.

[b]LDL-C: low-density lipoprotein cholesterol.

[c]LDL-C level was calculated using the Friedewald equation [37].

[d]HDL-C: high-density lipoprotein cholesterol.

[e]c-PDS: clinical Portfolio Diet Score.

## Dietary Adherence to the Portfolio Diet

Table 2 shows the dietary adherence to the Portfolio Diet for the full score (c-PDS), which ranges from 0 to 25 points, and the 5 individual components, which range from 0 to 5 points. The primary outcome of dietary adherence to the Portfolio Diet increased by 1.25 (SD 2.8; 1.25/25, 5%) points in the app group (*P*=.28) and 0.19 (SD 4.4; 0.19/25, 1%) points in the control group (*P*=.93), although neither increase was statistically significant (*P*=.62) from baseline and there was no difference between groups. On the basis of our sample size, the effect size (1.06), and the pooled SD (3.69), the estimated power to detect a statistically significant between-group difference was 7.8% (1-β) with an =.05, so due to the sample size, we were underpowered to detect a significant difference in dietary adherence between groups.

**Table 2.** Dietary adherence to the Portfolio Diet from weighed 7-day diet records, measured using the clinical Portfolio Diet Score (week 0 to week 12)[a].

| | App group (n=7) | | | | Control group (n=5) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Week, mean (SD) | | Δ[b], mean (SD) | *P* value[c] | Week, mean (SD) | | Δ, mean (SD) | *P* value[c] | *P* value[d] |
| | 0 | 12 | | | 0 | 12 | | | |
| Nuts and seeds, points | 3.4 (1.2) | 3.6 (1.6) | 0.2 (1.8) | .82 | 2.8 (1.2) | 2.7 (1.6) | –0.1 (2.3) | .92 | .82 |
| Plant protein, points | 2.8 (1.1) | 2.6 (1.3) | –0.2 (0.8) | .54 | 3.2 (1.7) | 2.9 (2.3) | –0.3 (0.7) | .48 | .91 |
| Viscous fiber, points | 3.3 (1.5) | 2.8 (1.8) | –0.5 (0.9) | .21 | 2.6 (1.7) | 2.2 (1.2) | –0.4 (0.9) | .32 | .94 |
| Plant sterols, points | 2.0 (1.8) | 3.6 (1.8) | 1.6 (1.9) | .08 | 3.5 (1.7) | 4.3 (0.5) | 0.7 (1.8) | .41 | .46 |
| High MUFA[e] oils and foods, points | 1.6 (1.1) | 1.8 (1.2) | 0.2 (0.9) | .56 | 1.6 (1.8) | 1.9 (1.6) | 0.3 (0.8) | .48 | .91 |
| Total c-PDS[f], points | 13.2 (3.7) | 14.5 (5.1) | 1.3 (2.8) | .28 | 13.7 (5.8) | 13.9 (5.2) | 0.2 (4.4) | .93 | .62 |

[a]The individual components are shown in points (range 0 to 5), which make up the total c-PDS (range 0 to 25).

[b]Δ represents change.

[c]*P* value for within group.

[d]*P* value for across groups.

[e]MUFA: monounsaturated fatty acid.

[f]cPDS: clinical Portfolio Diet Score.

## Acceptability

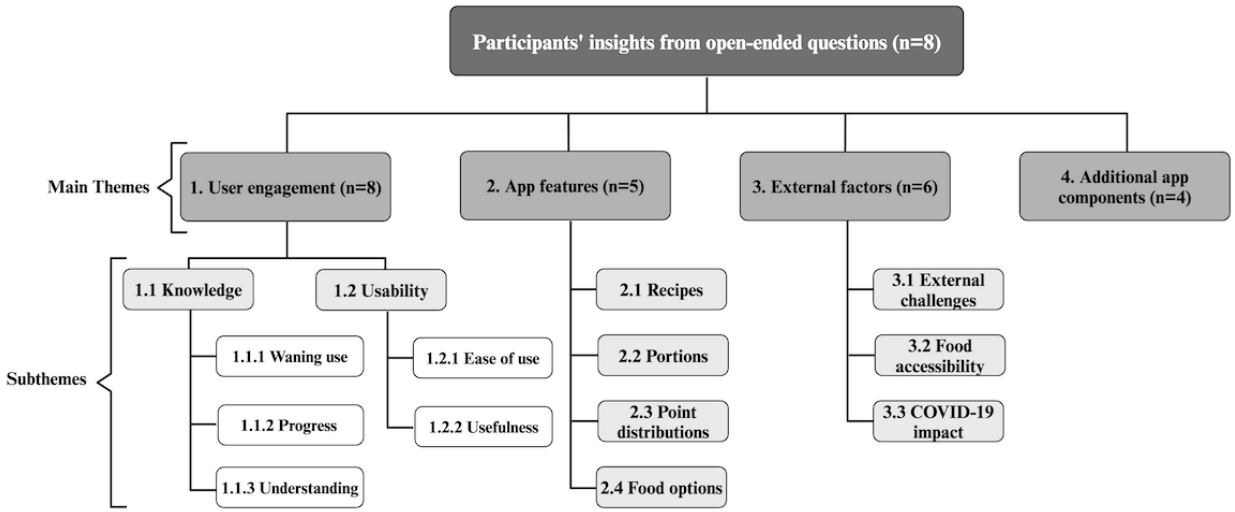
Multimedia Appendix 4 shows the average PortfolioDiet.app use by intervention month. Participants logged into the app an average of 18 (SD 14) days per month over the 12-week intervention period with the number of log-ins trending down over the duration of the intervention but these results were not statistically significant (Table S1 in Multimedia Appendix 5). The average SUS score was 80.9 (SD 17.3), which surpasses the usability quality benchmark threshold of 70, indicating a high level of usability [35]. Table S2 in Multimedia Appendix 5 shows the scores for individual SUS items. The individual responses to the SUS items (range 1-5) show that most participants felt confident using the app (mean 4.0, SD 1.31), they thought the app was easy to use (mean 4.25, SD 1.16), and they felt that the various functions in the PortfolioDiet.app were well linked together (mean 4.5, SD 0.76). Table S3 in Multimedia Appendix 5 summarizes the quantitative responses to the questionnaire**.** More than half of the participants (5/8, 63%) agreed that using the app increased their knowledge about the Portfolio Diet. Tip sheets and email reminders were ranked as the top app features for helping participants learn about the diet and support their interest or engagement in using the app, respectively.

## Participant Insights From Open-Ended Questions

### *Overview*

Figure 3 presents the results of the qualitative data assessments of open-ended questions. The open-ended questions expanded upon the SUS, providing contextual insights into participants' responses. A total of 4 main themes were identified: user engagement, app features, external factors, and added components. Each theme was further categorized into subthemes. Table S4 in Multimedia Appendix 5 presents individual participant quotations categorized under these themes related to their experience using the PortfolioDiet.app. Notably, 1 participant's insights were excluded from the table as their questionnaire responses were retrieved through a telephone conversation, wherein a member of the research team documented the responses. However, the insights provided by this participant were considered during data analysis.

**Figure 3.** Overview of main themes and subthemes identified from open-ended question responses. The number of participants with statements in each main theme is indicated by "(n=)".




## *Theme 1: User Engagement*

### Overview

The theme user engagement describes participants' experiences using the PortfolioDiet.app and sheds light on how they actively used, responded to, and integrated the app into their lives. Within this overarching theme, we found that participants described their engagement in various ways that could be divided into two subthemes: (1) knowledge, relating to participants' knowledge acquisition on the Portfolio Diet, which was further subdivided into waning use, progress, and understanding, and (2) usability, relating to the usability of the PortfolioDiet.app, which was further subdivided into usefulness and ease of use.

### Knowledge

#### *Waning Use*

It relates to how participants' engagement with the app transformed over time, revealing a pattern of gradual decline. Some participants mentioned that as they became more acquainted with the Portfolio Diet principles, their initial enthusiasm diminished. This sentiment of diminished engagement appeared to be rooted in the perception that the app's educational value was more pronounced during the early stages of app use:

> *I think the app is for new users. After you get up to speed and figure out how to do the [Portfolio Diet] and how [to] split your portions throughout the day, I can't see using the app daily for me.* [Participant 6]

#### *Progress*

Most participants acknowledged the app's role in helping them learn about their progress on the Portfolio Diet. Some participants referenced the leaderboard feature as being insightful in tracking their progress and understanding where their Portfolio Diet Score (PDS) stands. One participant expressed that the tracking or progress monitoring feature of the PortfolioDiet.app provided them with a sense of being actively engaged in their progress:

> *I enjoy tracking as it keeps me on target for food intake.* [Participant 3]

Another participant mentioned that the leaderboard encouraged them to "cheat more rather than eat more [Portfolio Diet] foods." However, other participants appreciated the app's tracking and progress monitoring features as they contributed to a sense of accountability and competition, motivating participants to align their dietary choices with the Portfolio Diet principles.

#### *Understanding*

Participants commented on how the app enriched their comprehension of the Portfolio Diet. Some participants articulated how the app's clear instructions and visual aids enhanced their understanding of the diet. One participant emphasized the ingenuity of the app's concept and its thoughtful design:

> *I think the concept is very clever and built in a meaningful way.... I have a much better understanding of the diet and how I am supposed to follow it.* [Participant 5]

### Usability

#### *Ease of Use*

When exploring the app's usability, participants elaborated on their impressions of the app's user friendliness. Participants largely found the app intuitive and easy to use. One participant noted that they had been following the Portfolio Diet for 3 years before incorporating the app into their routine. They found that using the app for tracking purposes was more convenient and practical compared to using a traditional paper checklist:

> *I was already on the third year of the Portfolio Diet when I started using the app. For me, it was easier [and] more handy to track using the app than using a checklist on paper.* [Participant 4]

Others mentioned a learning curve associated with using the app, noting the transition from requiring assistance to gaining confidence in using the app:

> *I was somewhat worried about the complexity of the app but got over it after the first couple of days of trying it out.* [Participant 5]

Other participants echoed a similar sentiment in their feedback regarding uncertainties about specific aspects of the PortfolioDiet.app. For instance, 1 participant provided positive feedback about the weekly questions for points, but voiced confusion over the meaning of star points and their implications:

> *The weekly questions for points were an interesting addition that I liked. I could not figure out what the star points meant when I logged out. I couldn't find an explanation if you miss a certain number of days or a certain threshold of daily points that you would slide backwards in the 30 day points graph.* [Participant 6]

Over a telephone interview, a participant also highlighted their concern with some technical features of the app, mentioning that the responsiveness of the bars within the app was slower than desired and reporting occasional log-in issues.

#### *Usefulness*

The usefulness of the PortfolioDiet.app was described by participants when evaluating the app's usability in their daily routines. A participant shared that the app offered them a unique perspective by focusing on helpful ways to enhance their PDS. By incorporating advice from the PortfolioDiet.app into their routine, they were able to make actionable behavior changes. As described by this participant, adding the liquid plant sterol supplement to their breakfast routine was an easy and impactful way to increase their PDS by 5 points:

> *[The app] helped me look at how to increase my daily [Portfolio Diet] score. For example, after I started using the app, I got into a regular use of the [plant] sterol supplement with my oatmeal every morning. My use of these supplements was more sporadic but using the app made me appreciate the high value of the supplement.* [Participant 5]

Alternatively, other participants mentioned they felt that the PortfolioDiet.app did not provide any additional incentives beyond their regular one-on-one meetings with trained dietitians, as part of the ongoing PortfolioEx trial:

> *There was nothing more in the app than what we were taught to do.* [Participant 2]

### *Theme 2: App Features*

After reviewing the feedback provided by participants, it became evident that several features of the PortfolioDiet.app were prominently mentioned. Specifically, participants emphasized the recipes, portions, point distribution (PDS), and food options.

#### Recipes

Notably, regarding recipes, one participant found them enjoyable to try, while another appreciated the app's inclusion of recipes but did not find that they aligned with their eating style. One participant described the recipes as a "nice addition" but mentioned that they did not try any of them:

> *The recipes were a nice addition however, I am a simple eater and didn't try any of the recipes. It is difficult to assess how one of my recipes or a vegan recipe book could be converted so I just assume if it has lots of oat bran or soy within, then it fits with the Portfolio diet.* [Participant 6]

Conversely, a different participant provided constructive feedback, suggesting that a review of the recipes might be beneficial, as they noted instances where certain ingredients or complete instructions were missing.

#### Portions

The participant feedback encompassed a range of viewpoints regarding the portion sizes recommended by the PortfolioDiet.app. While 1 participant described the portion sizes as "helpful," others voiced concerns that they appeared "enormous," "confusing," and seemingly tailored for a "higher calorie diet":

> *Initially, the app portion sizes were confusing .... Some portions on the app (i.e., barley) appeared enormous and put me off.* [Participant 6]

Two participants drew comparisons between the traditional paper checklist from the PortfolioEx trial they used to track their adherence to the Portfolio Diet and the app's portion feature, detailing the hurdles they encountered during the learning process. In addition, they emphasized discrepancies between the app's portion feature and their accustomed checklist. One of the participants described the following:

> *I did not like that it didn't line up exactly with the Daily checklist sheets which I used for about a year or more and got used to the portions and amounts on these sheets. It didn't line up. I also didn't like at first that I couldn't change it to my caloric intake.* [Participant 7]

#### Point Distribution (PDS)

Participants commented on how the point distribution component of the PortfolioDiet.app enabled them to monitor their scores, identify if they were high or low, and explore opportunities to improve their scores through changing aspects of their diet in accordance with the Portfolio Diet principles. One participant described the following:

> *The app was most helpful in delineating the different categories and how to improve your score if you were low in one of the five categories.* [Participant 5]

Alternatively, the same participant described how the organization of the point distribution components "frustrated" them as they did not align with the portion sizes they usually ate, evident in the following statement:

> *However, I found myself to be a little frustrated in some of the way the points are distributed. Using the viscous fibre category as an example that [highlights] the frustration I manage to eat at least an orange or an apple a day but not 2. Also, I eat a fair bit of ... eggplant but never 4 cups worth in one sitting.* [Participant 5]

Furthermore, another participant shared their experience of confusion while calculating points, expressing uncertainty about the value of food servings in terms of points:

> *At times, it is confusing calculating points. An example is the Oils. For 1 tsp of oil is the point "1" or "2" points?* [Participant 3]

#### Food Options

The feedback received consisted mainly of participant approval of the selection of foods included in the PortfolioDiet.app. However, 1 participant articulated desiring a broader range of food choices in the PortfolioDiet.app:

> *I hope one day the app can be used to track more foods to the categories.* [Participant 5]

In addition, another participant expressed contentment with the app's food options, attributed to the convenience of locating these items at the grocery store.

### *Theme 3: External Factors*

On the basis of the analysis of the participant's feedback from the intervention arm, external factors were identified as one of the main themes. External factors explored influences mentioned by participants that either positively or negatively impacted their ability to follow the Portfolio Diet but were not related to the app.

#### External Challenges

Participants mentioned some barriers in following the Portfolio Diet that were not directly related to the app or the COVID-19 pandemic. One participant expressed that the act of traveling posed challenges in adhering to the Portfolio Diet recommendations. While not elaborated upon, this sentiment highlights the real-world implications of dietary interventions, where external factors such as travel can impact the ability to follow dietary interventions:

> *Travelling makes it more difficult to follow [the Portfolio Diet].* [Participant 2]

A different participant expressed experiencing fatigue from adhering to the intervention. The participant's remark indicates that maintaining adherence to the Portfolio Diet can become challenging over time. This insight underscores the potential external factors, such as lack of novelty, that can influence an individual's engagement with this dietary intervention:

> *It's me getting tired of following a vegan diet.* [Participant 4]

#### Food Accessibility

Comments on the practicality of accessing recommended foods for the Portfolio Diet were captured as an important area for understanding how the Portfolio Diet can be applied to diverse populations. A participant shared that they use soy foods and shelf-stable soy milk from a particular store, likely due to the convenience it offered. They also mentioned finding an alternative plant sterol powder at a specific store, which they incorporated into their diet. This account provides valuable insight into the participant's resourcefulness in adapting their dietary habits to the Portfolio Diet, especially when faced with challenges like limited availability of certain products:

> *I find soy foods in the freezer aisle of Loblaws and use the shelf life Soy milk so I don't have to go to the store so often during Covid... I found a [plant] sterol powder at Healthy Planet that substitutes for the [plant] sterol margarine that's no longer produced and it's good in shakes or in my all-bran buds cereal ....* [Participant 6]

While the only comment made in this study about food accessibility was positive, we emphasize future work on the Portfolio Diet to capture future participants' feedback on this subtheme.

#### COVID-19 Pandemic Impact

As this study was run during the COVID-19 pandemic, a specific open-ended question related to its impact on the participants was included within the questionnaire. Understanding how participants from various situations experienced the COVID-19 pandemic and how it impacted their adherence to the Portfolio Diet may influence interpretation of the results of the study. Participants mentioned issues related to a lack of in-person meetings with the study dietitians and gym closures, while others articulated how they had been self-sufficient and were able to find study foods independently outside of the clinic. Interestingly, as the study was at the "tail end" of the lockdown, the impact of business reopening was noted by 1 participant:

> *Yes, with lock down, I was able to follow the diet very well, but since opening up, I have been more inclined to eat out and also crave foods that I haven't had in a long time at my favorite restaurants.... Definitely have felt some slow down in my incentive to keep strictly to the diet since the reopening. Also we are travelling a bit and I am excited to try the foods of the region we are travelling in so I also strayed from the Portfolio regime as a result.* [Participant 7]

### *Theme 4: Added App Components*

Participants articulated suggestions for app improvements and several requests, including the ability to record half portions, more food suggestions, visual meal plans, and more information related to diabetes. A participant pointed out the app's lack of capability for personalized adjustments to their dietary plan, which the dietitians had been able to offer them individually. This feedback underscores the value of personalized guidance and highlights a potential area for improvement in the app's functionality to better accommodate individualized dietary adjustments:

> *The app doesn't allow for personal [tweaking] to the portfolio as the dietitians have been able to do for me personally.* [Participant 7]

Some recommendations for features were already embedded within the app. As an example, 1 user suggested including the option to record half portions of food, a feature already available on the PortfolioDiet.app. This feedback indicates that the participant was not aware of this feature, suggesting it was not intuitive. Overall, we found that there were no overlapping

suggestions from participants, demonstrating the importance of ensuring the app can be personalized to any user based on their needs and preferences.

## Discussion

### Principal Findings

We conducted a 12-week randomized controlled ancillary mixed methods study to assess the effect of the PortfolioDiet.app on dietary adherence and its acceptability among high-risk adults. Although adherence was higher for the PortfolioDiet.app group after 12 weeks (ie, increased by 1.25/25, 5% and 0.19/25, 1% in the app group and control group, respectively), no difference between the groups was observed in this small ancillary study.

The PortfolioDiet.app was rated as usable, with the app surpassing the usability quality benchmark threshold [35]. While participants engaged often with the app over the 12 weeks, use gradually declined. Beyond the usability, the app increased self-reported knowledge of the Portfolio Diet. The demonstration of increased knowledge in those who had already been learning about the Portfolio Diet for an average of approximately 2 years further supports the acceptability of the app in this high-risk population. These results shed light on the potential of app-based technology as a promising platform to translate the Portfolio Diet to adults at high CVD risk.

The decline in use combined with the trending increase in adherence to the Portfolio Diet from 7DDRs, aligns with the intended purpose of the app as an educational tool aimed at fostering users' self-efficacy. As participants become more knowledgeable and confident in applying the principles of the Portfolio Diet, it is expected that their reliance on the app and use of the tracking progress feature would gradually decrease. However, based on participant feedback, modifications to the app to make this expectation clear to the user may further improve app acceptability. This messaging could include a note on the role the app can play for users at various times in their life, when they perhaps fall off the diet and need support to return to following the Portfolio Diet.

The qualitative data assessments complemented the quantitative findings. Analysis of open-ended questions identified 4 primary themes that encapsulated participants interactions with the PortfolioDiet.app. Among the themes, "user engagement" underscores the dynamic interactions participants had with the app, their knowledge gained, and the integration of its features into their routine. This was also evident in the quantitative findings which revealed that most participants felt that various functions of the PortfolioDiet.app were well linked together. The app's usefulness for self-monitoring of dietary adherence was noted as important and helpful by some participants. The educational aspect of the app was a recurrent point of mention among participants, with several of them noting how it enhanced or aided their current understanding of the Portfolio Diet. This observation aligns with the quantitative finding where more than half of the participants said that the app increased their knowledge of the Portfolio Diet. On the other hand, comments suggesting that the app provided no new information beyond what was provided in their regular one-on-one meetings with trained dietitians may provide an indication of why others may have responded "No" to this question about increasing knowledge on the Portfolio Diet. As all participants had already been participating in the PortfolioEx trial learning about the Portfolio Diet, this finding suggests the app is reinforcing counseling from dietitians.

The second theme, "app features," highlighted features participants found helpful or frustrating. These findings align with the current understanding as self-tracking and gamification features have been found as successful tools in health apps for behavior change [41]. However, some features of the app, such as the portions, could be better explained by using pop-up windows with additional instructions or through other modifications to the app.

The theme "external factors" delved into influences beyond the app's control on dietary adherence. Notably, the impact of the COVID-19 pandemic was explored, revealing its implications on participants' adherence patterns as pandemic restrictions shifted.

The fourth theme "additional app components" covered participants' feedback to include additional features to the app. Participants expressed a desire for additional food options and visual meal plans, as well as more diabetes-related information. Other desirable app modifications can be distilled from comments relating to the dislike of certain features (eg, leaderboard), challenges in logging foods, and adding half portion sizes. These comments imply possible modifications to the app that could improve its usability and acceptability, such as features of the app that need to be more intuitive and the ability for users to customize their own targets and dashboard.

Identifying that tip sheets and videos supported learning and engagement in the app can be leveraged in addressing some of the challenges identified by participants. Tip sheets could be developed to include tips while traveling or on the go, for meal plan ideas, and further support for those with diabetes. Integrating an interactive frame within the app to showcase new content, such as tip sheets, as well as videos to further support engagement may be a useful modification based on the participant feedback. Taken together, these findings suggest that the PortfolioDiet.app has the potential to support participants in adhering to the Portfolio Diet and is considered acceptable by adults at high CVD risk.

### Comparison With Prior Work

This study is the first to use the PortfolioDiet.app in high-risk adults. While health apps have seen widespread adoption, findings have been inconsistent when looking at their effects on behavior change and health outcomes. Similar to our findings, a systematic review and meta-analysis of 47 studies revealed that web-based interventions targeting risk factors show promise in reducing CVD risk, yet their effects were moderate and waned over time [42]. Inconsistencies in effects may be related to differences in the app features, the participant's health status, and whether the app intervention has been tailored to the population.

Apps that target dietary behavior change have also shown promise with suggestion that in those with chronic disease, use

of health apps with nutrition components improved health outcomes, with 64% of studies showing sustained behavior change for 6 to 12 months [43]. These conclusions differ from others who found health benefits were only observed in short-term studies (less than 6 months), suggesting that secondary prevention participants may be more motivated to make sustained behavior change.

When looking at health apps focused on delivering a therapeutic dietary pattern, a systematic review of 5 studies in participants with hypertension or prehypertension, found that mobile apps providing the Dietary Approaches to Stop Hypertension diet were associated with higher adherence to this diet and lower blood pressure when compared to controls [44]. However, the authors could not pinpoint the most effective features of these apps from a users' perspective. Identifying specific features may not be entirely possible as different population groups may prefer different strategies [43], emphasizing the importance of tailoring health apps to their intended population and allowing for personalization within the app. Interestingly, qualitative analysis of other health apps have identified similar themes with "new features" being identified as 1 of the 3 themes in adolescences with knee pain [45], mirroring our theme "Added app components." Without specific prompts, this shared interest underscores a patient's desire to shape tools meant to assist them and the importance of involving them in the cocreation process.

Several qualitative studies have identified barriers to nutrition app use. König et al [46] found that app usability was important for sustained uptake. The PortfolioDiet.app has been deemed usable in both a convenience sample of users and in our current representative sample of participants. When comparing our usability score to others in the literature, a raw SUS score of 80 would be better than 75% of all apps tested; however, average SUS scores varies based on the type of app being tested [47]. A systematic review of health apps found an average SUS score of 76.6 (SD 15.12), but when excluding physical activity apps, the average SUS dropped to 68.1 (SD 14.05) [48]. This finding aligns with the general understanding that nutrition apps are challenged with usability issues [46]. Specific to nutrition, an analysis of the top 7 diet-tracking apps (from iOS iTunes and Android Play web-based stores) found an average SUS of 70.9 (SD 12.72) with a range from 46.7 to 89.2, after 3 undergraduate nutrition students used the apps over a 2-week period [49].

In addition, personalized and tailored educational material, reminders, progress tracking, and goal setting have been found to be highly valued features [50], all of which are present in the PortfolioDiet.app. The usability and knowledge acquisition demonstrated in this study also aligns with the results of a previous quality assessment study of the PortfolioDiet.app in a convenience sample of users [26].

### Strengths and Limitations

The primary strength of this study is the assessment of the PortfolioDiet.app within its intended target population of adults at high risk of CVD, allowing for modifications to the app to support its use in the intended users. The collection of both quantitative and qualitative data is also a strength of this study as it allowed for a comprehensive understanding of participants' experiences with the PortfolioDiet.app. In addition, the synergy between the SUS findings along with the insights derived from qualitative analysis, where participants largely found the app intuitive and easy to use, strengthens our confidence that the app was considered usable by this study population. The influence of the COVID-19 pandemic on participants' experiences and engagement underscores the significance of remote health care solutions in ensuring quality care delivery despite challenging circumstances.

A major limitation was the restricted pool of participants, exacerbated by delays in the REB review due to the COVID-19 pandemic, among other challenges experienced by the research community [51]. These challenges led to a sample below the estimated necessary sample size, with the estimated power to detect a statistically significant between-group difference being 7.8% (1–β), =.05, so we were underpowered to detect a significant difference in dietary adherence between groups. The limited sample size should also be considered when interpreting the qualitative findings. While data saturation may be achievable with relatively small samples (9-17 interviews) [39], our sample falls below this range, so a cautious interpretation of the results is necessary.

In addition, we did not measure health-related risk factors directly. While much of the research in the realm of health apps has shown improvements in behaviors, there remains a notable gap in the literature concerning their impact on intermediate risk factors and other health outcomes. Consequently, it is imperative that future research endeavors incorporate assessments of health outcomes, such as lipid profiles, to provide a more comprehensive understanding of the impact of these apps on health and disease outcomes.

In addition, in light of research findings suggesting that marginalized populations may also experience digital exclusion exacerbating existing health disparities, it is crucial to emphasize the necessity of future research involving underserved groups [52].

Finally, the use of the SUS is another limitation as it was not originally tailored for evaluating health apps. However, the 100-point scale facilitates clear communication to nonexperts in the field. Moreover, the concise nature of the SUS, featuring 10 questions, ensures swift participant completion and reduces response burden, which is especially important when participants are not visiting the study center and instead are completing the questionnaires remotely. Possibly related to its high ease of use, the SUS was used in 40 of the 96 studies in a scoping review of health apps in older (>65 years) individuals [53]. Although other questionnaires to assess the usability of mobile health (mHealth) apps have recently been developed, the SUS remains widely used and considered suitable for assessing digital health apps [48,54]. However, to enhance specificity to mHealth apps, future evaluations of the PortfolioDiet.app administering questionnaires could include the user-oriented Mobile Application Rating Scale or the recently validated mHealth App Usability Questionnaire, which includes additional questions to integrate feedback on app features [55,56].

## Implications and Future Directions

As CVD continues to be a leading cause of mortality in Canada and globally [57], prioritizing lifestyle interventions for disease prevention and management is pivotal. Among these interventions, the Portfolio Diet is an effective therapy for managing dyslipidemia and reducing the risk of CVD. As a tool for disseminating this nutrition therapy, the PortfolioDiet.app may serve to increase the adoption of the Portfolio Diet.

Notably, there is growing interest among older adults in using mobile apps to support their learning efforts. In a survey conducted among Canadian retired older adults (aged >55 years), 78.5% agreed or strongly agreed that mobile devices made their learning easier [58], highlighting the potential of the PortfolioDiet.app to engage and empower older individuals, who are a critical demographic for cardiovascular health management. This observation underscores the substantial implications of the PortfolioDiet.app and the importance of tailoring the app to ensure older adults can engage with the app. From this study, we can discern both the app's strengths and limitations in its intended population of high-risk adults. These insights will guide us in refining the PortfolioDiet.app, creating a tool that better meets the needs of its target population. Subsequent work will incorporate the feedback received through modification to the design of the PortfolioDiet.app. While this work was undertaken in older high-risk adults, further research is needed in more diverse and underserved populations.

## Conclusions

This small ancillary study suggests the PortfolioDiet.app is considered acceptable, easy to use, and increases knowledge of the Portfolio Diet in adults at high CVD risk. The present findings highlight the potential of the PortfolioDiet.app as an educational tool, reinforcing counseling from dietitians. In general, participants appreciated the app's self-monitoring features as they contributed to a sense of accountability, motivating participants to align their dietary choices with the Portfolio Diet principles. Future refinements to ensure the app is intuitive and its features are well explained and can be personalized could enhance participant engagement and adherence to the Portfolio Diet for improved cardiovascular health. We await the results of a randomized controlled trial investigating the effect of the PortfolioDiet.app on lipid targets in a high-risk population, which may provide evidence of its potential health benefits.

## Acknowledgments

This research was funded by the Canadian Institutes of Health Research (CIHR) Doctoral Research Award (FRN: 181403). The Diet, Digestive tract, and Disease (3D) Centre, funded through the Canada Foundation for Innovation and the Ministry of Research and Innovation's Ontario Research Fund, provided the infrastructure for the conduct of this work. MEK was supported by the CIHR Doctoral Research Award FRN: 181403 and a Toronto 3D Knowledge Synthesis and Clinical Trials foundation PhD Scholarship Award. LC was a Mitacs-Elevate postdoctoral fellow jointly funded by the government of Canada and the Canadian Sugar Institute (September 2019-August 2021) and a Toronto 3D Knowledge Synthesis and Clinical Trials foundation New Investigator Award. AJG was supported by a CIHR Postdoctoral Fellowship. SA-C was funded by a CIHR Canadian Graduate Scholarships Master's Award, the Loblaw Food as Medicine Graduate Award, the Ontario Graduate Scholarship, and the CIHR Canadian Graduate Scholarship Doctoral Award 476251. AZ was funded by a Toronto 3D Knowledge Synthesis and Clinical Trails Unit foundation fellowship. DJAJ was funded by the Government of Canada through the Canada Research Chair Endowment. JLS was funded by a PSI Graham Farquharson Knowledge Translation Fellowship; Canadian Diabetes Association Clinician Scientist Award; Canadian Institute of Health Research INMD and CNS New Investigator Partnership Prize; and Banting and Best Diabetes Centre Sun Life Financial New Investigator Award. VSM was supported by the Canada Research Chairs Program; Connaught New Researcher Award, University of Toronto; The Joannah & Brian Lawson Centre for Child Nutrition, University of Toronto; and Temerty Faculty of Medicine Pathway Grant, University of Toronto.

The authors would like to thank all participants for their time and detailed feedback. The authors would also like to thank their PortfolioDiet.app team and the many volunteers who assisted in the development of the PortfolioDiet.app [59].

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

MEK, LC, and JLS were responsible for conceptualization. MEK, LC, SMQ, KR, NA, MP, SS-P, DP, SMG, AJG, SA-C, AZ, RGJ, VSM, CWCK, DJAJ, and JLS were responsible for methodology and writing—review and editing for important intellectual content. MEK, KR, NA, MP, SS-P, and DP were responsible for data collection. MEK, LC, SMQ, and GV were responsible for qualitative analysis. MEK was responsible for statistical analysis and writing—original draft preparation. JLS was responsible for supervision. MEK and JLS were responsible for funding acquisition. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

## Conflicts of Interest

MEK was a part-time employee at INQUIS Clinical Research, Ltd, a contract research organization. LC has received research support from the Canadian Institutes of Health Research (CIHR), Protein Industries Canada (a government of Canada Global Innovation Clusters), Alberta Pulse Growers, and the United Soybean Board (USDA soy Checkoff program). AJG has received travel support and/or honoraria from Vinasoy, the Soy Nutrition Institute Global, and the Academy of Nutrition and Dietetics. SA-C has received an honorarium from the International Food Information Council for a talk on artificial sweeteners, the gut microbiome, and the risk for diabetes. AZ is a part-time research associate at INQUIS Clinical Research, Ltd, a contract research organization. She has received consulting fees from the Glycemic Index Foundation. CWCK, DJAJ, and JLS have received funding support, honoraria, consulting, or travel fees from a broad range of food, beverage, and ingredient companies, trade associations, government agencies, health charities, private foundations, or other commercial or nonprofit entities with an interest in nutrition and chronic disease prevention and management. For a complete list of disclosures, see Multimedia Appendix 6. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
CONSORT 2010 checklist.
[PDF File (Adobe PDF File), 87 KB - cardio_v9i1e58124_app1.pdf ]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1743 KB - cardio_v9i1e58124_app2.pdf ]

Multimedia Appendix 3
PortfolioDiet.app participant feedback questionnaire.
[DOCX File , 22 KB - cardio_v9i1e58124_app3.docx ]

Multimedia Appendix 4
Average days logged into the PortfolioDiet.app over the intervention (12 weeks; n=8).
[PNG File , 62 KB - cardio_v9i1e58124_app4.png ]

Multimedia Appendix 5
Supplemental tables including use, usability, and feedback on the PortfolioDiet.app.
[DOCX File , 37 KB - cardio_v9i1e58124_app5.docx ]

Multimedia Appendix 6
Full list of all disclosures.
[DOCX File , 27 KB - cardio_v9i1e58124_app6.docx ]

## References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020 Dec 22;76(25):2982-3021 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
2. Sud M, Han L, Koh M, Abdel-Qadir H, Austin PC, Farkouh ME, et al. Low-density lipoprotein cholesterol and adverse cardiovascular events after percutaneous coronary intervention. J Am Coll Cardiol 2020 Sep 22;76(12):1440-1450 [FREE Full text] [doi: 10.1016/j.jacc.2020.07.033] [Medline: 32943162]
3. Sarak B, Savu A, Kaul P, McAlister FA, Welsh RC, Yan AT, et al. Lipid testing, lipid-modifying therapy, and PCSK9 (proprotein convertase subtilisin-Kexin Type 9) inhibitor eligibility in 27 979 patients with incident acute coronary syndrome. Circ Cardiovasc Qual Outcomes 2021 Apr;14(4):e006646. [doi: 10.1161/circoutcomes.120.006646]
4. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. Eur Heart J 2017 Aug 21;38(32):2459-2472 [FREE Full text] [doi: 10.1093/eurheartj/ehx144] [Medline: 28444290]
5. Desai NR, Farbaniec M, Karalis DG. Nonadherence to lipid-lowering therapy and strategies to improve adherence in patients with atherosclerotic cardiovascular disease. Clin Cardiol 2023 Jan;46(1):13-21 [FREE Full text] [doi: 10.1002/clc.23935] [Medline: 36267039]

6. Lichtenstein AH, Appel LJ, Vadiveloo M, Hu F, Kris-Etherton P, Rebholz C, et al. 2021 dietary guidance to improve cardiovascular health: a scientific statement from the American Heart Association. Circulation 2021 Dec 07;144(23):e472-e487 [FREE Full text] [doi: 10.1161/CIR.0000000000001031] [Medline: 34724806]
7. Anderson TJ, Grégoire J, Pearson GJ, Barry AR, Couture P, Dawes M, et al. 2016 Canadian cardiovascular society guidelines for the management of dyslipidemia for the prevention of cardiovascular disease in the adult. Can J Cardiol 2016 Nov;32(11):1263-1282. [doi: 10.1016/j.cjca.2016.07.510] [Medline: 27712954]
8. Pearson GJ, Thanassoulis G, Anderson TJ, Barry AR, Couture P, Dayan N, et al. 2021 Canadian cardiovascular society guidelines for the management of dyslipidemia for the prevention of cardiovascular disease in adults. Can J Cardiol 2021 Aug;37(8):1129-1150. [doi: 10.1016/j.cjca.2021.03.016] [Medline: 33781847]
9. Diabetes Canada Clinical Practice Guidelines Expert Committee, Punthakee Z, Goldenberg R, Katz P. Definition, classification and diagnosis of diabetes, prediabetes and metabolic syndrome. Can J Diabetes 2018 Apr;42 Suppl 1:S10-S15. [doi: 10.1016/j.jcjd.2017.10.003] [Medline: 29650080]
10. Brown J, Clarke C, Johnson SC, Sievenpiper J. Canadian adult obesity clinical practice guidelines: medical nutrition therapy in obesity management. Obesity Canada. URL: https://obesitycanada.ca/guidelines/nutrition [accessed 2021-03-25]
11. Tobe SW, Stone JA, Anderson T, Bacon S, Cheng AY, Daskalopoulou SS, et al. Canadian Cardiovascular Harmonized National Guidelines Endeavour (C-CHANGE) guideline for the prevention and management of cardiovascular disease in primary care: 2018 update. CMAJ 2018 Oct 09;190(40):E1192-E1206 [FREE Full text] [doi: 10.1503/cmaj.180194] [Medline: 30301743]
12. Portfolio diet. Heart UK: The Cholesterol Charity. 2016. URL: https://www.heartuk.org.uk/downloads/healthprofessionals/factsheets/the-portfolio-diet.pdf [accessed 2016-02-04]
13. Stroes ES, Thompson PD, Corsini A, Vladutiu GD, Raal FJ, Ray KK, European Atherosclerosis Society Consensus Panel. Statin-associated muscle symptoms: impact on statin therapy-European Atherosclerosis Society Consensus Panel Statement on Assessment, Aetiology and Management. Eur Heart J 2015 May 01;36(17):1012-1022 [FREE Full text] [doi: 10.1093/eurheartj/ehv043] [Medline: 25694464]
14. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines. J Am Coll Cardiol 2019 Jun 25;73(24):3168-3209 [FREE Full text] [doi: 10.1016/j.jacc.2018.11.002] [Medline: 30423391]
15. Jenkins DJ, Kendall CW, Marchie A, Faulkner DA, Wong JM, de Souza R, et al. Effects of a dietary portfolio of cholesterol-lowering foods vs lovastatin on serum lipids and C-reactive protein. JAMA 2003 Jul 23;290(4):502-510. [doi: 10.1001/jama.290.4.502] [Medline: 12876093]
16. Chiavaroli L, Nishi SK, Khan TA, Braunstein CR, Glenn AJ, Mejia SB, et al. Portfolio dietary pattern and cardiovascular disease: a systematic review and meta-analysis of controlled trials. Prog Cardiovasc Dis 2018;61(1):43-53 [FREE Full text] [doi: 10.1016/j.pcad.2018.05.004] [Medline: 29807048]
17. Kris-Etherton PM, Akabas SR, Bales CW, Bistrian B, Braun L, Edwards MS, et al. The need to advance nutrition education in the training of health care professionals and recommended research to evaluate implementation and effectiveness. Am J Clin Nutr 2014 May;99(5 Suppl):1153S-1166S [FREE Full text] [doi: 10.3945/ajcn.113.073502] [Medline: 24717343]
18. Aboueid S, Bourgeault I, Giroux I. Nutrition and obesity care in multidisciplinary primary care settings in Ontario, Canada: short duration of visits and complex health problems perceived as barriers. Prev Med Rep 2018 Jun;10:242-247 [FREE Full text] [doi: 10.1016/j.pmedr.2018.04.003] [Medline: 29868375]
19. Wynn K, Trudeau JD, Taunton K, Gowans M, Scott I. Nutrition in primary care: current practices, attitudes, and barriers. Can Fam Physician 2010 Mar;56(3):e109-e116 [FREE Full text] [Medline: 20228290]
20. Brauer PM, Sergeant LA, Davidson B, Goy R, Dietrich L. Patient reports of lifestyle advice in primary care. Can J Diet Pract Res 2012;73(3):122-127. [doi: 10.3148/73.3.2012.122] [Medline: 22958629]
21. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: a systematic review of randomized controlled trials. Healthc Inform Res 2018 Jul;24(3):207-226 [FREE Full text] [doi: 10.4258/hir.2018.24.3.207] [Medline: 30109154]
22. Burn E, Nghiem S, Jan S, Redfern J, Rodgers A, Thiagalingam A, et al. Cost-effectiveness of a text message programme for the prevention of recurrent cardiovascular events. Heart 2017 Jun;103(12):893-894 [FREE Full text] [doi: 10.1136/heartjnl-2016-310195] [Medline: 28235776]
23. Cheung KL, Wijnen BF, Hiligsmann M, Coyle K, Coyle D, Pokhrel S, et al. Is it cost‐effective to provide internet‐based interventions to complement the current provision of smoking cessation services in the Netherlands? An analysis based on the EQUIPTMOD. Addiction 2017 Dec 15;113(S1):87-95. [doi: 10.1111/add.14069]
24. Bates DW, Landman A, Levine DM. Health apps and health policy: what is needed? JAMA 2018 Nov 20;320(19):1975-1976. [doi: 10.1001/jama.2018.14378] [Medline: 30326025]
25. The portfolio diet. PortfolioDiet.app. URL: https://portfoliodiet.app/ [accessed 2021-11-03]

26. Kavanagh ME, Chiavaroli L, Glenn AJ, Heijmans G, Grant SM, Chow CM, et al. A web-based health application to translate nutrition therapy for cardiovascular risk reduction in primary care (portfoliodiet.app): quality improvement and usability testing study. JMIR Hum Factors 2022 Apr 21;9(2):e34704 [FREE Full text] [doi: 10.2196/34704] [Medline: 35451981]
27. Aschbrenner KA, Kruse G, Gallo JJ, Plano Clark VL. Applying mixed methods to pilot feasibility studies to inform intervention trials. Pilot Feasibility Stud 2022 Sep 26;8(1):217. [doi: 10.1186/s40814-022-01178-x]
28. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
29. Creswell JW, Clark VL. Designing and Conducting Mixed Methods Research. Thousand Oaks, CA: Sage Publications; 2007.
30. Create a blocked randomisation list. Sealed Envelope Ltd. URL: https://www.sealedenvelope.com/simple-randomiser/v1/lists/ [accessed 2021-03-08]
31. Nishi SK, Kavanagh ME, Ramboanga K, Ayoub-Charette S, Modol S, Dias G, et al. Effect of digital health applications with or without gamification on physical activity and cardiometabolic risk factors: a systematic review and meta-analysis of randomized controlled trials. eClinicalMedicine 2024 Oct;76:102798. [doi: 10.1016/j.eclinm.2024.102798]
32. Kavanagh ME, Glenn AJ, Chiavaroli L, Morgan GA, Josse RG, Malik VS, et al. Simulation model to assess the validity of the clinical portfolio diet score used in the PortfolioDiet.app for dietary self-tracking: a secondary analysis of a randomized controlled trial in hyperlipidemic adults. Front Nutr 2024 Aug 7;11:1398450 [FREE Full text] [doi: 10.3389/fnut.2024.1398450] [Medline: 39171108]
33. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London, UK: CRC Press; 1996:189-194.
34. Lewis JR, Sauro J. The factor structure of the system usability scale. In: Proceedings of the 1st International Conference, HCD 2009, Held as Part of HCI International on Human Centered Design. 2009 Presented at: HCD '09; July 19-24, 2009; San Diego, CA p. 94-103 URL: https://link.springer.com/chapter/10.1007/978-3-642-02806-9_12 [doi: 10.1007/978-3-642-02806-9_12]
35. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
36. M Grant S, J Glenn A, M S Wolever T, G Josse R, L O'Connor D, Thompson A, et al. Evaluation of glycemic index education in people living with type 2 diabetes: participant satisfaction, knowledge uptake, and application. Nutrients 2020 Aug 12;12(8):25 [FREE Full text] [doi: 10.3390/nu12082416] [Medline: 32806563]
37. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972 Jun;18(6):499-502. [Medline: 4337382]
38. Jenkins DJ, Jones PJ, Lamarche B, Kendall CW, Faulkner D, Cermakova L, et al. Effect of a dietary portfolio of cholesterol-lowering foods given at 2 levels of intensity of dietary advice on serum lipids in hyperlipidemia: a randomized controlled trial. JAMA 2011 Aug 24;306(8):831-839. [doi: 10.1001/jama.2011.1202] [Medline: 21862744]
39. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
40. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology: Research Designs: Quantitative, Qualitative, Neuropsychological, Biological, Vol 2. 2nd edition. Washington, DC: American Psychological Association; 2023:65-81.
41. Davis AJ, Parker HM, Gallagher R. Gamified applications for secondary prevention in patients with high cardiovascular disease risk: a systematic review of effectiveness and acceptability. J Clin Nurs 2021 Oct;30(19-20):3001-3010. [doi: 10.1111/jocn.15808] [Medline: 33872436]
42. Beishuizen CR, Stephan BC, van Gool WA, Brayne C, Peters RJ, Andrieu S, et al. Web-based interventions targeting cardiovascular risk factors in middle-aged and older people: a systematic review and meta-analysis. J Med Internet Res 2016 Mar 11;18(3):e55 [FREE Full text] [doi: 10.2196/jmir.5218] [Medline: 26968879]
43. Salas-Groves E, Galyean S, Alcorn M, Childress A. Behavior change effectiveness using nutrition apps in people with chronic diseases: scoping review. JMIR Mhealth Uhealth 2023 Jan 13;11:e41235 [FREE Full text] [doi: 10.2196/41235] [Medline: 36637888]
44. Alnooh G, Alessa T, Hawley M, de Witte L. The Use of Dietary Approaches to Stop Hypertension (DASH) mobile apps for supporting a healthy diet and controlling hypertension in adults: systematic review. JMIR Cardio 2022 Nov 02;6(2):e35876 [FREE Full text] [doi: 10.2196/35876] [Medline: 36322108]
45. Johansen SK, Kanstrup AM, Thomsen JL, Christensen MN, Rathleff MS. Exploring the barriers and facilitators for supporting adolescents with knee pains adherence to mobile health apps: a think-aloud study. Digit Health 2023 Oct 19;9:20552076231205750 [FREE Full text] [doi: 10.1177/20552076231205750] [Medline: 37868153]
46. König LM, Attig C, Franke T, Renner B. Barriers to and facilitators for using nutrition apps: systematic review and conceptual framework. JMIR Mhealth Uhealth 2021 Apr 01;9(6):e20037 [FREE Full text] [doi: 10.2196/20037] [Medline: 34254938]
47. Sauro J. A Practical Guide to Measuring Usability: 72 Answers to the Most Common Questions about Quantifying the Usability of Websites and Software. Denver, CO: Measuring Usability LLC; 2010.

48. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
49. Ferrara G, Kim J, Lin S, Hua J, Seto E. A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. JMIR Mhealth Uhealth 2019 May 17;7(5):e9232 [FREE Full text] [doi: 10.2196/mhealth.9232] [Medline: 31102369]
50. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158. [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
51. Sathian B, Asim M, Banerjee I, Pizarro AB, Roy B, van Teijlingen ER, et al. Impact of COVID-19 on clinical trials and clinical research: a systematic review. Nepal J Epidemiol 2020 Sep;10(3):878-887 [FREE Full text] [doi: 10.3126/nje.v10i3.31622] [Medline: 33042591]
52. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. J Med Internet Res 2020 Jun 02;22(6):e19361. [doi: 10.2196/19361] [Medline: 32452816]
53. Wang Q, Liu J, Zhou L, Tian J, Chen X, Zhang W, et al. Usability evaluation of mHealth apps for elderly individuals: a scoping review. BMC Med Inform Decis Mak 2022 Dec 02;22(1):317 [FREE Full text] [doi: 10.1186/s12911-022-02064-5] [Medline: 36461017]
54. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022 Jan 27;22(1):22 [FREE Full text] [doi: 10.1186/s12911-022-01764-2] [Medline: 35081953]
55. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
56. Zhou L, Bao J, Setiawan IM, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
57. Report from the Canadian Chronic Disease Surveillance System: heart disease in Canada. Public Health Agency of Canada. 2018. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/report-heart-disease-Canada-2018.html [accessed 2024-04-29]
58. Morrison D, Koole M. Learning on-the-go: older adults? Use of mobile devices to enhance self-directed, informal learning. J Interact Learn Res 2018;29(3):423-443.
59. Our team: the portfolio diet app. Portfolio Diet Recipes. URL: https://recipes.portfoliodiet.app/our-team/ [accessed 2021-09-10]


## Abbreviations

**7DDR:** 7-day diet record
**CONSORT:** Consolidated Standards of Reporting Trials
**c-PDS:** clinical Portfolio Diet Score
**CVD:** cardiovascular disease
**LDL-C:** low-density lipoprotein cholesterol
**mHealth:** mobile health
**PDS:** Portfolio Diet Score
**REB:** research ethics board
**SUS:** System Usability Scale


*Edited by A Coristine; submitted 06.03.24; peer-reviewed by K Mendoza, J Mistry, J Alfonsi; comments to author 14.11.24; revised version received 09.01.25; accepted 03.02.25; published 28.03.25.*

*Please cite as:*
*Kavanagh ME, Chiavaroli L, Quibrantar SM, Viscardi G, Ramboanga K, Amlin N, Paquette M, Sahye-Pudaruth S, Patel D, Grant SM, Glenn AJ, Ayoub-Charette S, Zurbau A, Josse RG, Malik VS, Kendall CWC, Jenkins DJA, Sievenpiper JL*
*Acceptability of a Web-Based Health App (PortfolioDiet.app) to Translate a Nutrition Therapy for Cardiovascular Disease in High-Risk Adults: Mixed Methods Randomized Ancillary Pilot Study*
*JMIR Cardio 2025;9:e58124*
*URL: https://cardio.jmir.org/2025/1/e58124*
*doi:10.2196/58124*
*PMID:*

Original Paper

# Co-Occurring Diseases and Mortality in Patients With Chronic Heart Disease, Modeling Their Dynamically Expanding Disease Portfolios: Nationwide Register Study

Nikolaj Normann Holm[1], PhD; Anne Frølich[2,3], PhD; Helena Dominguez[4], PhD; Kim Peder Dalhoff[5,6], PhD; Helle Gybel Juul-Larsen[7], PhD; Ove Andersen[6,7,8], PhD; Anders Stockmarr[1], PhD

[1]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark

[2]Innovation and Research Centre for Multimorbidity, Slagelse Hospital, Slagelse, Denmark

[3]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[4]Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark

[5]Department of Clinical Pharmacology, Bispebjerg and Frederiksberg University Hospital, Copenhagen, Denmark

[6]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

[7]Department of Clinical Research, Copenhagen University Hospital Amager and Hvidovre, Hvidovre, Denmark

[8]Emergency Department, Copenhagen University Hospital Amager and Hvidovre, Hvidovre, Denmark

**Corresponding Author:**
Nikolaj Normann Holm, PhD
Department of Applied Mathematics and Computer Science
Technical University of Denmark
Richard Petersens Plads, Building 324
Kgs. Lyngby, 2800
Denmark
Phone: 45 45 25 30 31
Email: nnho@dtu.dk

## Abstract

**Background:** Medical advances in managing patients with chronic heart disease (HD) permit the co-occurrence of other chronic diseases due to increased longevity, causing them to become multimorbid. Previous research on the effect of co-occurring diseases on mortality among patients with HD often considers disease counts or clusters at HD diagnosis, overlooking the dynamics of patients' disease portfolios over time, where new chronic diseases are diagnosed before death. Furthermore, these studies do not consider interactions among diseases and between diseases, biological and socioeconomic variables, which are essential for addressing health disparities among patients with HD. Therefore, a mapping of the effect of combinations of these co-occurring diseases on mortality among patients with HD considering such interactions in a dynamic setting is warranted.

**Objective:** This study aimed to examine the effect of the co-occurring diseases of patients with HD on mortality, modeling their dynamically expanding chronic disease portfolios while identifying interactions between the co-occurring diseases, socioeconomic and biological variables.

**Methods:** This study used data from the national Danish registries and algorithmic diagnoses of 15 chronic diseases to obtain a study population of all 766,596 adult patients with HD in Denmark from January 1, 1995, to December 31, 2015. The time from HD diagnosis until death was modeled using an extended Cox model involving chronic diseases and their interactions as time-varying covariates. We identified interactions between co-occurring diseases, socioeconomic and biological variables in a data-driven manner using a hierarchical forward-backward selection procedure and stability selection. We mapped the impact on mortality of (1) the most common disease portfolios, (2) the disease portfolios subject to the highest level of interaction, and (3) the most severe disease portfolios. Estimates from interaction-based models were compared to an additive model.

**Results:** Cancer had the highest impact on mortality (hazard ratio=6.72 for male individuals and 7.59 for female individuals). Excluding cancer revealed schizophrenia and dementia as those with the highest mortality impact (top 5 hazard ratios in the 11.72-13.37 range for male individuals and 13.86-16.65 for female individuals for combinations of 4 diseases). The additive model underestimated the effects up to a factor of 1.4 compared to the interaction model. Stroke, osteoporosis, chronic obstructive

pulmonary disease, dementia, and depression were identified as chronic diseases involved in the most complex interactions, which were of the fifth order.

**Conclusions:** The findings of this study emphasize the importance of identifying and modeling disease interactions to gain a comprehensive understanding of mortality risk in patients with HD. This study illustrated that complex interactions are widespread among the co-occurring chronic diseases of patients with HD. Failing to account for these interactions can lead to an oversimplified attribution of risk to individual diseases, which may, in cases of multiple co-occurring diseases, result in an underestimation of mortality risk.

***(JMIR Cardio 2025;9:e57749)*** doi:10.2196/57749

**KEYWORDS**

survival analysis; interaction effects; chronic heart disease; multimorbidity; time-varying covariates

# Introduction

## Background

Driven by the advancements in diagnostic tools and medical treatments, the mortality of patients with chronic heart disease (HD) has decreased considerably [1]. However, with a prolonged life span comes a risk of developing additional chronic diseases and complications to their HD [2], causing them to become multimorbid [3]. Multimorbidity is highly prevalent among patients with HD [2,4,5], and the increasing disease burden may modify time to death [6].

Recent research has identified the most prevalent comorbidities in patients with HD and how they affect mortality and other adverse health-related outcomes [5,7-9]. However, only a few studies have considered the effect of several diseases in the same person. Among these studies, there is a large variety in which diagnoses are considered and which statistical methods are applied. The studies that consider multimorbidity either restrict their analyses to a subset of diagnosis combinations [7] or group diagnoses into multimorbidity clusters at baseline before analyzing the effects of the extracted clusters [5]. Despite modeling disease interactions, these kinds of analyses fail to capture the crucial dynamics in the HD disease trajectories, where additional diseases are cumulatively diagnosed before death [10], causing an augmented risk profile for the patient. As the chronology of disease onset has been associated with a change in mortality among common diagnoses [11], it is thus essential to consider this dynamic development when analyzing effects. Due to the high prevalence of multimorbidity among patients with HD, the unique combination of chronic diseases that a patient has at any given time—referred to as the *disease portfolio*—is not static. Instead, it evolves over the observation period as new chronic diseases develop. This dynamic expansion reflects the progressive accumulation of chronic diseases in an individual following their HD diagnosis until death. As only a few studies consider these dynamics, there is a need for a thorough, large-scale study of the impact of disease interactions on mortality, modeling such a dynamic expansion of the patients' disease portfolios. Such an investigation would enable obtaining a deeper understanding of how the complexity of disease progression in patients with HD affects mortality over time.

The significance of understanding the effects of the emergence of multimorbidity over an individual's life span has previously been highlighted [3,12,13]. However, rather than treating multimorbidity as a singular risk factor, we took a more nuanced approach by dissecting the effects of multimorbidity based on the diseases appearing in the disease portfolio, recognizing that each combination of chronic diseases can affect mortality differently. Furthermore, as many chronic diseases have similar biological and socioeconomic risk factors, knowledge of the interplay between the impact of these is essential and can be used for possible preventive interventions and the development of guidelines for relevant coexisting diseases [14,15]. For instance, consider a disease portfolio comprising HD and osteoporosis. The impact on the mortality hazard rate may vary between men and women. Expanding on this example, the effect of socioeconomic position may differ depending on both sex and the presence of osteoporosis in the portfolio. These variations in effects represent interactions in modeling terms. As such, identifying and emphasizing interactions between chronic diseases and demographic factors can shed new light on the impact of pathophysiological pathways on mortality.

## Objectives

This large-scale study is based on data from the total adult Danish population recorded in nationwide primary and secondary health care registries, including medical diagnoses, medications, educational attainment level, and health care use. We used an extended Cox model with time-varying covariates to model time until death for individuals diagnosed with HD considering their dynamically expanding disease portfolios. In our model, the hazard ratio (HR) of a disease portfolio is constant. In contrast, the HR of an individual changes dynamically when the individual obtains a new portfolio by developing a new chronic disease (Figure 1).

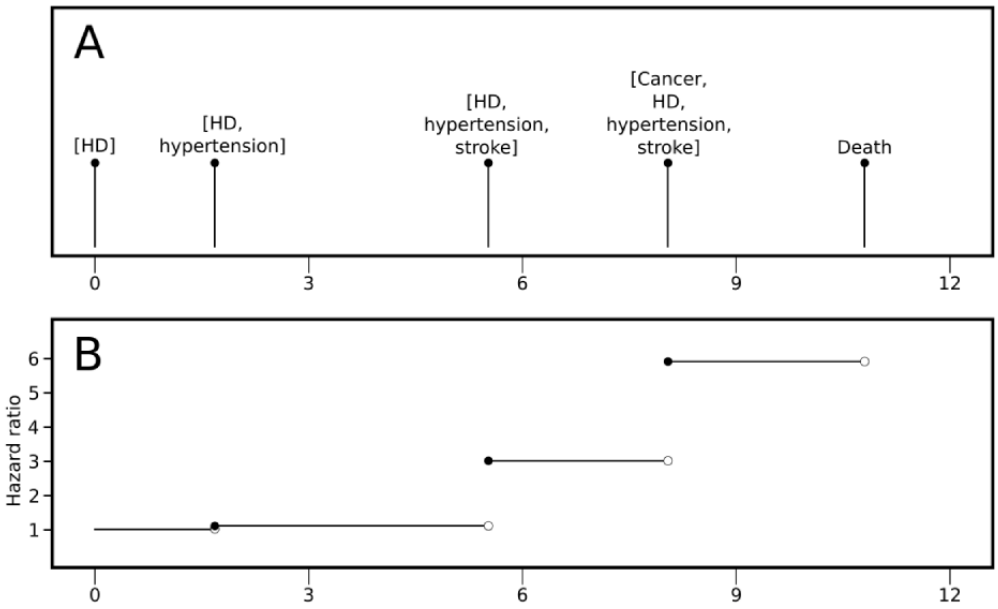
We conducted a model and data-driven selection of interaction effects. Subsequently, we studied the impact on time to death according to the (1) most frequently occurring disease portfolios, (2) most complex disease portfolios in terms of order of interactions, and (3) disease portfolios with the highest hazards relative to only HD.

We recognize the inherent complexity in interpreting interaction effects, especially in the case of higher-order interactions involving multiple factors. However, to emphasize the importance of modeling interaction effects, we also present a comparative analysis of effect estimates for disease portfolios, contrasting our interaction model with a simpler model in which interactions are excluded. The differences observed in these

comparisons serve to underscore the crucial role of modeling interactions in medical research.

Throughout this paper, we use a bracket notation to represent the disease portfolio of a specific patient with HD. For example, a patient with HD, diabetes, and hypertension is denoted by the portfolio [diabetes, hypertension]. If the patient with HD also has high cholesterol, their disease portfolio is [diabetes, high cholesterol, hypertension]. As all individuals in the study population had HD, we use the term *disease portfolio* without mentioning the coexisting HD diagnosis in the notation. We use the terms *dyads*, *triads*, *tetrads*, and *pentads* to describe disease portfolios of size 2, 3, 4, and 5, respectively, with size being the number of chronic diseases in the portfolio including HD.

**Figure 1.** Example of how the statistical model works. (A) Illustration of an event sequence in which a hypothetical patient with heart disease (HD) receives the diagnosis of HD at time 0 and, subsequently, the hypertension, stroke, and cancer diagnoses at different times (measured in years following HD diagnosis) before death. (B) The corresponding longitudinal development of the hazard ratio of the patient relative to a theoretical patient who only has HD and is not multimorbid.




## Methods

### Data Foundation

All children born in Denmark or any new residents are, by law, required to obtain a unique personal identification number, which is stored in the Danish Civil Registration System [16]. The personal identification number can link information from any additional Danish register at an individual level subject to General Data Protection Regulation restrictions [17]. Information about chronic disease diagnoses was based on diagnostic algorithms initially developed by the Research Center for Prevention and Health at Glostrup University Hospital [18]. These algorithms cover 15 diagnoses based on their clinical relevance that have been previously used in national reports of chronic disease diagnoses in Denmark [19,20]. Moreover, previous work with these diagnoses has shown prevalence results comparable to those of other European studies [21]. The algorithmic diagnoses are based on data recorded in 4 registries: the Danish National Patient Register [22], the Danish Psychiatric Central Research Register [23], the Danish National Prescription Registry [24], and the Danish National Health Service Register [25]. Therefore, a particular diagnosis can be given at a particular time (with temporal granularity of days) based on criteria for hospitalization diagnoses, medication, or repeated use of specific health services. As such, a single diagnosis corresponds to 1 disease and represents multiple Anatomical Therapeutic Chemical or *ICD-10* (*International Statistical Classification of Diseases, 10th Revision*) codes with similar treatments and organization of health care. Thus, the diagnosis time stamps considered in this study are diagnostic time stamps and should not be regarded as time stamps for disease onset. In addition to the registries used for diagnostic time stamps, we used the Danish Population Education Register [26] and the Danish Register of Causes of Death [27] for information on educational attainment and death.

### Study Design and Population

Using our data foundation from the Danish registries, we obtained a study population of individuals diagnosed with HD covering the entire Danish adult population (aged ≥18 years) at some point during the observation period from January 1, 1995, to December 31, 2015, which had been previously analyzed [28]. These people were followed up on, and data associated with visits to outpatient clinics, hospital stays, primary sector health services, and prescriptions were collected for each person throughout the observation period. To define the study population, we applied algorithmic diagnoses (detailed in Multimedia Appendix 1) to identify individuals diagnosed with HD while determining diagnostic time stamps for 14 additional selected chronic diseases [21]. Thus, our inclusion criterion was broad, encompassing all Danish adults (aged ≥18 years) who received an algorithmic diagnosis of HD during the study period. No additional exclusion criteria were applied. Our

outcome was time until death of any cause after the HD diagnosis.

### Statistical Analysis

The prevalence of each of the chronic diseases was calculated at the time of HD diagnosis across all patients in the population. Similarly, we calculated prevalences of the diseases throughout the observation period by considering whether the condition had occurred at all among the patients with HD.

The data were analyzed within a survival analysis framework, with years following HD diagnosis as the time variable and an event defined as all-cause mortality. As such, we denoted the HD diagnosis time as $t$=0 and aligned our timescale accordingly, meaning that time $t$=0 corresponds to potentially different age times and calendar times for distinct individuals. In addition, individuals lost to follow-up due to emigration or reaching the end of the observation period were censored at these times.

The time-varying information on individual diagnoses; information on sex (male or female), age, educational attainment level (none, short, medium, long, missing, and missing before 1920); and calendar time were included as explanatory variables in the analysis (refer to tables 1/2 in the study by Holm et al [28]). We used an extended Cox model to estimate the effect of these explanatory variables on mortality, allowing for the inclusion of time-varying covariates. We classified our variables into primary and intrinsic categories [29]. Primary variables, such as the time-varying diagnosis indicators, cover variables of paramount interest. Intrinsic variables define the study individuals (ie, the variables sex, age, educational attainment level, and calendar time). Interactions both between and within each group of variables were considered. The numerical variables were mean centered before analysis.

As the development of additional diagnoses is a continuous process, the primary variables were allowed to change over time. These variables were piecewise constant in time, being 0 when the diagnosis was not present and 1 when obtained and onward in time. As the registries continuously cover clinical events for all individuals over the observation period, these diagnosis variables update at individual-specific time points dictated by the (sequence of) events that trigger the algorithmic diagnosis (Multimedia Appendix 1). An example of a potential sequence of diagnoses is showcased in Figure 1.

In the extended Cox proportional hazard model [30], the hazard $h_i$ for the $i$th individual at time $t$ is given by the following equation:

**(1)**

In this equation, $h_0(t)$ is the unspecified baseline hazard function for a male individual with no education without any diagnoses except HD. $X_{ij}(t)$ denotes the variable $j$ for individual $i$ (with $\boldsymbol{X}_i(t)$ denoting a vector of all variables) at time $t$, with $i$=1,...,$n$. The $\beta_j$ are the effect parameters. Due to $h_0(t)$ being unspecified, these parameters are linked to the relative mortality hazard rate of a variable as opposed to the absolute risk. Equation 1 assumes that variables have proportionate effects on the hazard function over time. We assessed this assumption for each variable by examining Schoenfeld residuals [31]. In addition, as the effect parameters $\beta_j$ do not depend on time, the hazard rate associated with a particular combination of explanatory variables was assumed to be the same across all time points.

To analyze the data, the following software was used: R (version 4.2.2; R Foundation for Statistical Computing), with the packages *survival* (version 3.5-5), *lava* (version 1.7.1), *glmnet* (version 4.1-6), and *multcomp* (version. 1_4-20).

### Selection of Variables and Interactions

It is essential to account for diagnosis interactions as such parameters serve to model the entire effect of disease portfolios associated with mortality. Possible omitted interaction effects from a model in which a significant interaction exists can result in a misrepresentation of the relationship between the variables and the time until death. It may also lead to bias in parameter estimation [32,33].

A common way to perform variable selection is a backward selection approach starting from a full model considering all possible interactions, reducing it to a model that best explains the observed data. However, such an approach was not computationally feasible as we are in a big data setting with numerous observations and countless potential variable interactions. Instead, we considered 2 variations of a forward-backward selection procedure to discover disease interactions. As a sensitivity analysis, we also performed variable selection using the stability selection methodology [34] with the regularization-based least absolute shrinkage and selection operator (LASSO) [35] approach.

In addition to the models including interaction effects, a model solely consisting of the primary and intrinsic variables' main effects (and squared and cubic terms) was estimated for reference.

We considered $k$-way interactions iteratively for $k$=2,..., $M$, with $M$ being a predetermined upper limit. The selection procedure starts from an initial model including all main effects and works in the following way for each value of $k$:

- Generate $n_c$ candidate variable additions obtained from the current model by adding a single $k$-way interaction to an already existing ($k$ – 1)–way interaction, also adding necessary lower hierarchical terms.
- Repeat until there are no candidate models below the cutoff: (1) estimate each of the candidate models obtained from adding any of the $n_c$ variables not already added to the current model and compare with the current model using a likelihood ratio test and (2) select the candidate model with the lowest $P$ value below the cutoff $\alpha/n_c$ as the current model.
- Clean up potentially masked significances in the $k$-way selection path through backward selection using a test level of $\alpha$.

The selection algorithm runs either until $M$-way interactions are included or until no $k$-way interactions are selected in the $k$th iteration. In the forward step of the selection algorithm, a

Bonferroni-adjusted cutoff is used to minimize the risk of false discoveries as each variable addition is potentially tested for inclusion $n_c$ times. We note that all considered models are hierarchical, meaning that, if a model contains a 5-way interaction among 5 variables, it also contains all possible 4-, 3-, and 2-way interactions among those variables.

Due to the allowance of *any* *k*-way interaction between and among the primary and intrinsic variables, a possibly large number of candidate models were included for each value of *k*. Because of this, the selection forward step was relaxed such that the candidate model *P* values were ordered from lowest to highest after the first estimation for each value of *k*. In the following estimations, variable additions were checked in this order, immediately adding any interactions below the cutoff while discarding insignificant terms. Before backward selection, any discarded terms were included again through forward selection. To introduce conservatism, all variable selections were performed with $\alpha$=.001. The resulting model with all selected interactions was labeled as the ALL model.

In addition to the ALL model, the variable selection procedure was run without relaxation of the forward step but only considering interactions among the primary variables. We labeled this as the disease interactions only (DIO) model. Furthermore, we used a variation of the stability selection framework [34], a method for improving variable selection in high-dimensional, sparse environments. This method selects variables repeatedly chosen on subsampled data through a structure learning method such as the LASSO algorithm for the Cox model [36]. We used a selection threshold of 0.9 following the recommendation in the work by Meinshausen and Bühlmann [34]. Each subsample included 10 randomly selected variables considering all their possible interactions up to an order of 5. This caused us to consider 3400 subsamples in total. We then fit an unregularized Cox model using the stably selected terms and performed backward selection to reduce the model using all available data. The resulting model was labeled the stable model. As a sensitivity criterion, we compared detected interactions among the chronic diseases across the ALL, DIO, and stable models. The additive model only including main effects was labeled as the only main effects (OME) model.

### Selecting Disease Portfolios

Due to the many possibilities when considering combinations of the 14 co-occurring diseases, some of our presented results are based on selected disease portfolios. These selections were made based on 3 criteria: most common disease portfolios, disease portfolios subject to the highest order of disease interactions, and disease portfolios with the highest mortality impact. The main results presented in this paper are based on the ALL model. To illustrate the importance of modeling interaction effects, the effect of specific disease portfolios in the ALL model was compared to additive effects from the OME model on the log-hazard scale.

### Scenarios

As the considered diagnosis variables were subject to higher-order interactions, effects were not apparent just from the estimated parameters because the effect of a single diagnosis varied across different levels of other diagnoses and intrinsic variables. To supplement the effect of the selected disease portfolios, the absolute mortality risk over time was estimated for multiple scenarios using the estimated ALL model. We did this to illustrate the modification of the risk profile over time of an individual diagnosed with HD. Each scenario represented the risk of a hypothetical individual whose disease portfolio expands at predetermined time points following HD diagnosis. The times at which the disease portfolio expanded in the hypothetical scenarios were determined in a data-driven fashion using gamma regressions, where the time points (at which the first, second, or third expansion of the disease portfolio following HD diagnosis occurred) were regressed on the diagnoses in the sequence considered in the scenario. The scenarios were constructed for patients who received their HD diagnosis at mean age and calendar time levels.

### Ethical Considerations

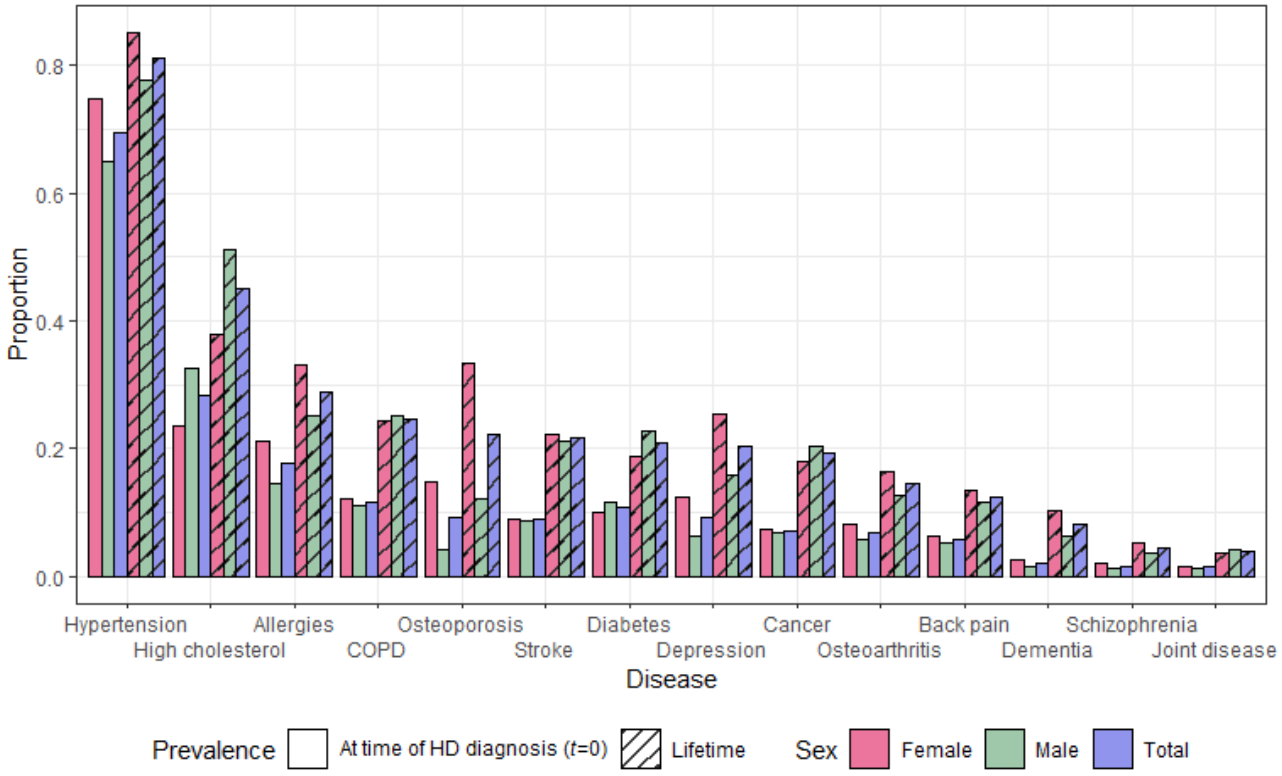
In this study, we used data from the national Danish registries, which are protected by the Danish Data Protection Act, meaning that they can only be accessed after application and subsequent approval. This study did not require additional approval from the Danish Research Ethics Committees or any informed consent as it solely involved the use of national registry data, exempt under the Scientific Ethical Committees Act. Danish registry data are deidentified to protect the privacy of individuals.

## Results

### Characteristics of the Study Population

A total of 766,596 individuals diagnosed with HD were included (n=406,792, 53.06% male). The mean age at the time of HD diagnosis was 67.51 (SD 13.07) years for male individuals and 73.02 (SD 13.37) years for female individuals (further baseline characteristics are available in table 2 in the work by Holm et al [28]). At the end of the observation period, 57.95% (444,233/766,596) were dead (222,112/406,792, 54.6% male and 222,121/359,804, 61.73% female). Overall, the prevalence of multimorbidity in the complete trajectories of each patient with HD was 96.88% (742,688/766,596). This was an increase compared to the multimorbidity prevalence at time *t*=0 (661,490/766,596, 86.29%). The prevalence of each of the 14 co-occurring diseases is presented in Figure 2. Overall, hypertension, high cholesterol, and allergies were among the most prevalent diseases in the HD population, with a lifetime prevalence of 81.18% (622,323/766,596), 44.94% (344,481/766,596), and 28.88% (221,385/766,596), respectively (Figure 2; Multimedia Appendix 2). Differences in prevalence by sex were large for some chronic diseases, particularly for osteoporosis and depression, commonly occurring in female individuals.

**Figure 2.** Diagnosis prevalence according to sex. Prevalence is reported at the time of heart disease (HD) diagnosis and for the entire span of the observed disease trajectories (Lifetime). COPD: chronic obstructive pulmonary disease.

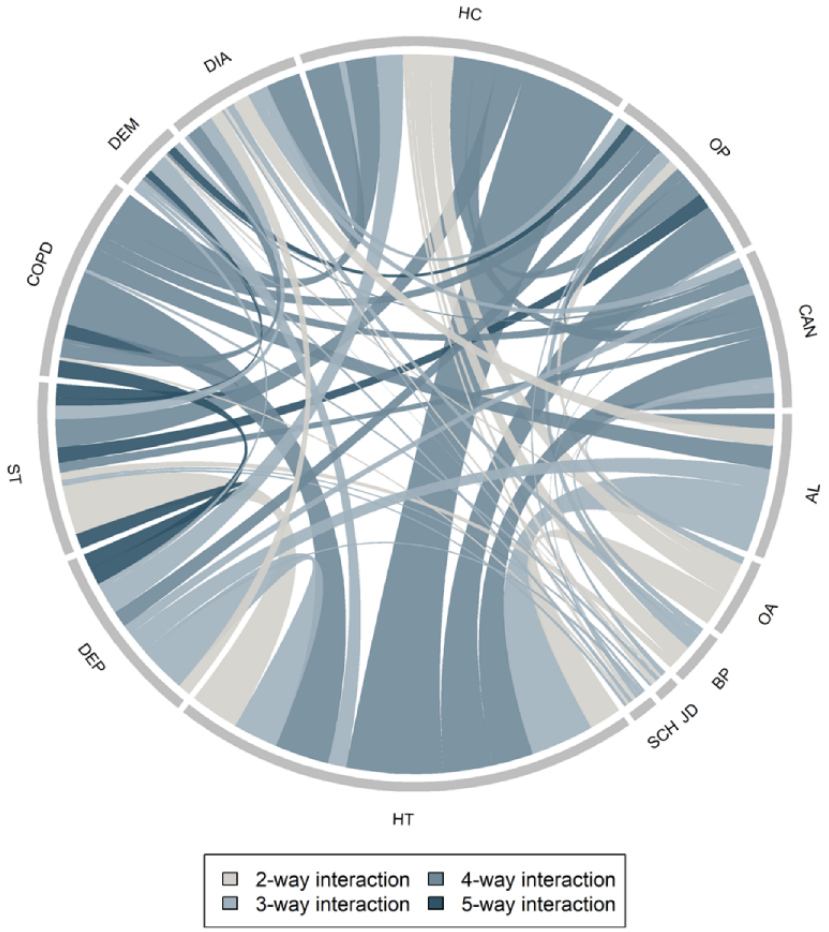



## Interactions

Following the inclusion of 5-way interactions, the ALL model selection procedure terminated due to no 6-way interactions being selected. All the primary and intrinsic variables were present in the final model. Figure 3 illustrates statistically significant ($P$<.001) interaction relationships between chronic diseases detected in the ALL model. Connections between diseases in the ribbon chart illustrate the 2 chronic diseases appearing in an interaction, with the color depicting the complexity of the interaction (darker color represents a higher-order interaction). The figure shows all diseases interacting, with some diseases involved in more complex interactions than other chronic diseases. In total, 288 interactions were present in the final model. The interaction relationships between the considered diseases were highly diverse but dominated by cancer, which had statistically significant interactions with all other diseases. Depression, stroke, chronic obstructive pulmonary disease (COPD), dementia, and osteoporosis were involved in the most complex interactions as they were the sole diseases involved in 5-way interactions. Some of the most prevalent diseases, allergies and hypertension, were not part of these complex relationships.

The chronic disease allergies were part of 5 interaction relationships with other diseases, involving two 4-way, two 3-way, and a single 2-way interaction. Hypertension interacted with 9 other diseases, involving four 4-way, three 3-way, and two 2-way interactions. Notably, dementia and depression appeared in higher-order interactions (two 5-way interactions) despite having fewer co-occurrences in the population. Similar patterns were observed for the DIO and stable models (Multimedia Appendices 3 and 4). In both models, COPD, dementia, stroke, and depression were involved in interactions of the highest order. The DIO model included up to 5-way interactions, also featuring complex interactions involving the chronic diseases diabetes and cancer (Multimedia Appendix 3). For the stable model, only up to 4-way interactions were detected (Multimedia Appendix 4). In general, most of the interactions between diseases identified in the ALL model were also present in the DIO and stable models (Multimedia Appendix 5).

**Figure 3.** Graphical representation of disease-disease interactions in the all interactions model. A ribbon connects chronic diseases that have any significant interaction (P<.001) between them. The connection's width corresponds to the number of individuals diagnosed with HD developing both diseases throughout the observation period. The ribbon's color represents the highest-order interaction relationship between 2 diseases. The ribbon chart is ordered by number of connections between diseases, starting from allergies (AL) with 5 connections all the way to cancer (CAN), which interacts with all the additional diseases. BP: back pain; COPD: chronic obstructive pulmonary disease; DEM: dementia; DEP: depression; DIA: diabetes; HC: high cholesterol; HT: hypertension; JD: joint disease; OA: osteoarthritis; OP: osteoporosis; SCH: schizophrenia; ST: stroke.

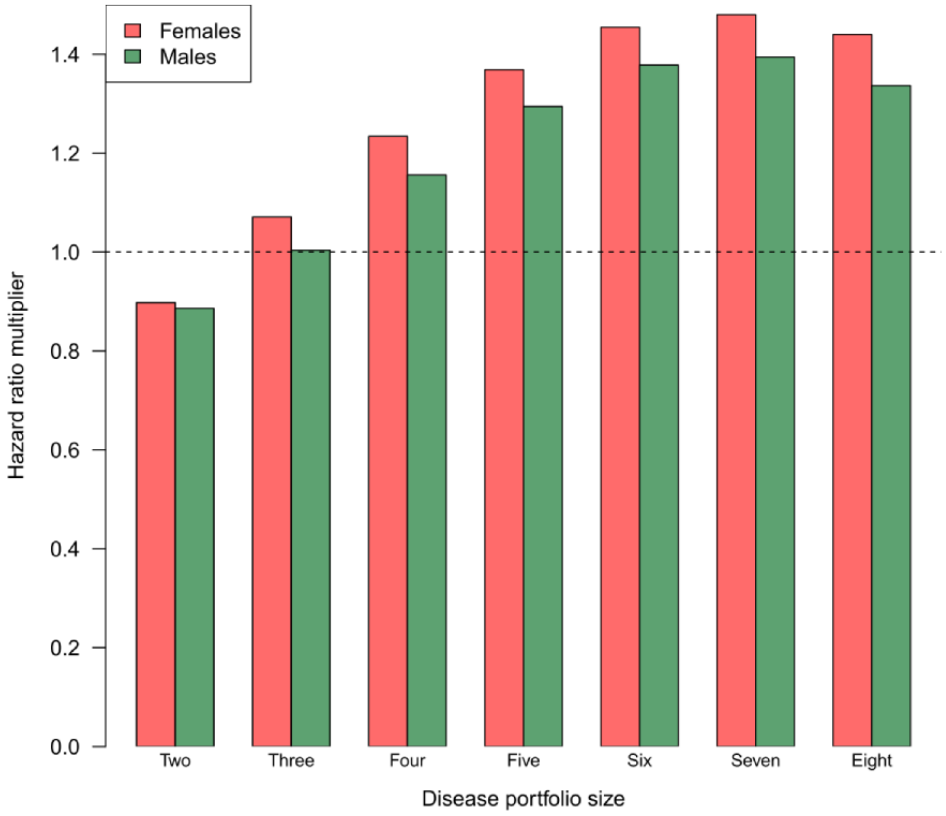



## Effects

### *Difference in Effect Estimates by Disease Portfolio Size*

To evaluate how the effects of disease portfolios on time until death differed between models with and without interactions, we calculated the effect differences between the OME model and the ALL model on the log-hazard rate scale, denoted as . We focused on disease portfolios ranging from 2 to 8 diseases as these accounted for 98.95% (1,671,575/1,689,297) of all disease portfolio observations of size ≥2. The effects in the ALL model for each disease portfolio were computed at mean age and calendar time levels, aggregating over combinations of both sexes and all educational attainment levels. To compute an overall estimate of the effect differences between the models for each sex, we calculated a weighted mean of the differences for each portfolio size. The weights were determined by the prevalence of individual disease portfolios across the different educational levels for each sex. In Figure 4, the aggregated differences are displayed on the hazard scale, indicating the multiplier required to convert the HR from the OME model into the HR from the ALL model. The figure illustrates substantial variations in disease portfolio effects when interactions were excluded compared to when they were included. The HR multiplier increased gradually for disease portfolios of increasing size, flattening at approximately 1.4 at disease portfolios of size 6. In general, for disease portfolios of size 2, the HRs were, on average, slightly overestimated when interactions were not modeled. However, for disease portfolios of size ≥4, the HRs were, on average, underestimated for both sexes. The underestimation also applies to female individuals with disease portfolios of size 3. In general, the HR multiplier was slightly greater for female individuals compared to male individuals across all disease portfolios.

**Figure 4.** Difference in effect estimates for disease portfolios of increasing size for female and male individuals. Each bar represents a weighted average of the differences in effects between the additive only main effects (OME) model and the all interactions (ALL) model on the hazard scale exp(Inline Graphic 3). Thus, the bars indicate the average multiplier required to convert the hazard ratio (HR) from the OME model into the HR from the ALL model. The weights were determined based on the occurrence of each specific disease portfolio across the different educational attainment levels for each sex.




### *Most Frequent Disease Portfolios*

The effects of the 10 most frequent disease portfolio dyads, triads, tetrads, and pentads are presented on the log-hazard scale at increasing educational attainment levels for male individuals in Figure 5 and for female individuals in Figure 6 based on the ALL model. The associated HR estimates are presented in Multimedia Appendices 6 and 7. Disease portfolios including high cholesterol and allergies were of particular concern as many of them had a negative effect, corresponding to a decreased mortality hazard rate relative to an individual diagnosed with HD who was not multimorbid. By comparing effects of the disease portfolios from the ALL model to effects from the OME model, generally, the direction of the effect (positive or negative) agreed between the models for both male and female individuals. However, the magnitude of the effects was greater in the ALL model than in the OME model for almost all disease portfolios, educational attainment levels, and sexes. This indicates an underestimation of the risk associated with a disease portfolio for the positive effects and an overestimation for the negative effects. For some disease portfolios, an inverse social gradient was visible in the educational dimension, where the higher the educational attainment level, the greater the effect of the disease portfolio (refer to, eg, the portfolio [diabetes, hypertension] in Figure 5). Sex-related disparities in disease portfolio effects were also evident. For disease portfolios containing depression and osteoporosis, the effects of the portfolios were greater for male individuals than for female individuals, whereas for COPD, cancer, stroke, and diabetes, the effects were greater for female individuals.

**Figure 5.** Effects of the 10 most frequent disease portfolio dyads (A), triads (B), tetrads (C), and pentads (D). Effects are shown for male individuals of varying educational attainment levels at the log-hazard rate scale. Comparisons are made to a male individual of the corresponding educational attainment level who only has heart disease (HD). Effects are presented for the all interactions model (different shades of blue) and the only main effects model (red). All comparisons are made at mean age and calendar time. HD is present in all disease portfolios. AL: allergies; BP: back pain; CAN: cancer; COPD: chronic obstructive pulmonary disease; DEP: depression; DIA: diabetes; HC: high cholesterol; HT: hypertension; OA: osteoarthritis; OP: osteoporosis; ST: stroke.

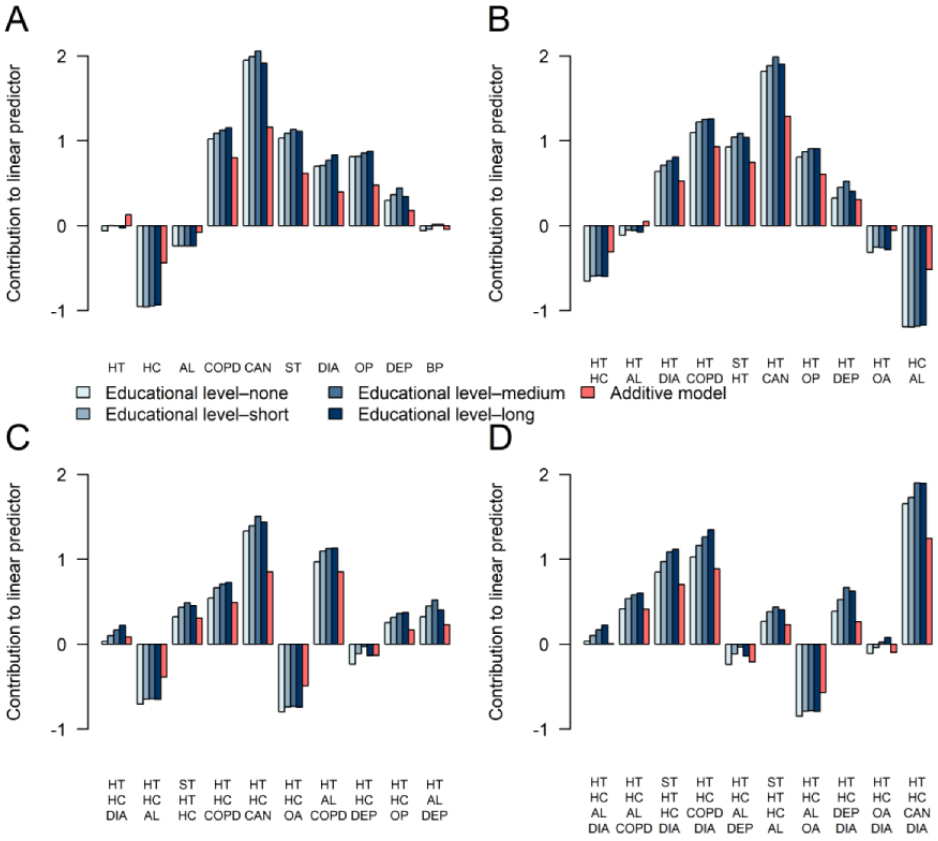
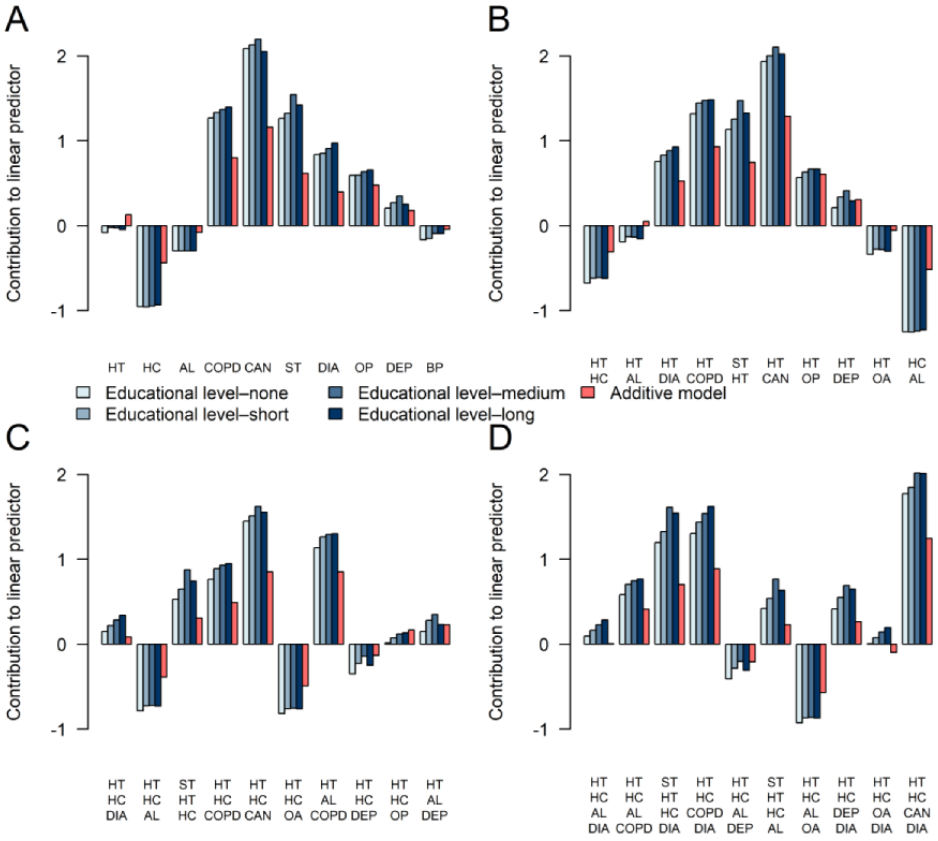


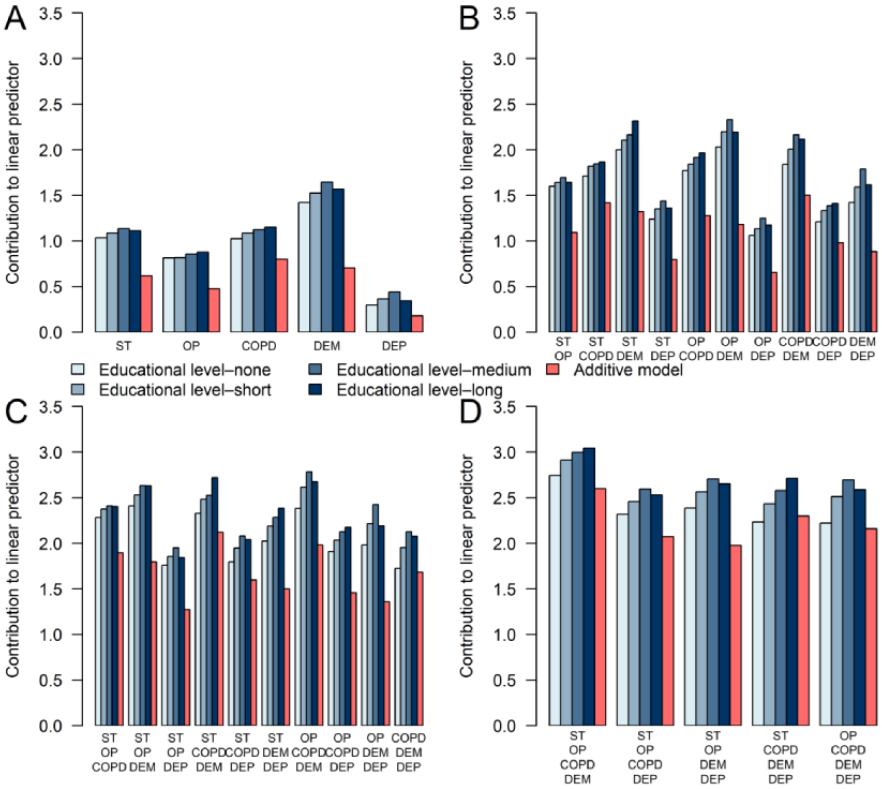
**Figure 6.** Effects of the 10 most frequent disease portfolio dyads (A), triads (B), tetrads (C), and pentads (D). Effects are shown for female individuals of varying educational attainment levels at the log-hazard rate scale. Comparisons are made to a female individual of the corresponding educational attainment level who only has heart disease (HD). Effects are presented for the all interactions model (different shades of blue) and the only main effects model (red). All comparisons are made at mean age and calendar time. HD is present in all disease portfolios. AL: allergies; BP: back pain; CAN: cancer; COPD: chronic obstructive pulmonary disease; DEP: depression; DIA: diabetes; HC: high cholesterol; HT: hypertension; OA: osteoarthritis; OP: osteoporosis; ST: stroke.




### *Most Complex Disease Portfolios*

Figure 7 shows the effects of disease portfolios containing combinations of stroke, osteoporosis, COPD, dementia, and depression for male individuals with differing educational attainment levels. These chronic diseases were all part of 5-way interactions, making the effects associated with their portfolios the most complex. For dyads, triads, tetrads, and pentads, the OME model generally yielded lower effects than the ALL model. This implies an underestimation of mortality risk in male individuals for these portfolios when interactions were not modeled. The underestimation was greatest for disease portfolios involving dementia or stroke. Similar results were observed for female individuals but also included a large underestimation of mortality hazard rates for portfolios involving COPD (Multimedia Appendix 8).

**Figure 7.** Effects of disease portfolio dyads (A), triads (B), tetrads (C), and pentads (D) involving stroke (ST), osteoporosis (OP), chronic obstructive pulmonary disease (COPD), dementia (DEM), and depression (DEP). Effects are shown for male individuals of varying educational attainment levels at the log-hazard rate scale. Comparisons are made to a male individual of the corresponding educational attainment level who only has heart disease (HD). Effects are presented for the all interactions model (different shades of blue) and the only main effects model (red). All comparisons are made at mean age and calendar time. HD is present in all disease portfolios.




### *Disease Portfolios With the Highest Mortality Impact*

Table 1 presents the largest HRs for disease portfolio dyads, triads, and tetrads among male and female individuals. Generally, the HRs of the disease portfolios were greater in female individuals; however, the portfolio [schizophrenia] exhibited a greater HR in male individuals. For dyads, the portfolios [cancer], [dementia], [schizophrenia], [stroke], and [COPD] ranked within the top 5 for both sexes. Notably, [cancer] exhibited the largest HR (6.72 for male individuals and 7.59 for female individuals). When considering triads and tetrads, cancer was similarly consistently featured in the top 5 portfolios for both sexes. This indicates that cancer contributes to a greatly increased relative mortality risk whenever present. Among triads, the portfolio [cancer, schizophrenia] had the largest HR for male individuals (13.26) and the second largest for female individuals (13.38). The top-ranking portfolio for female individuals was [cancer, COPD] (HR=15.39), whereas for male individuals, it was the second largest (HR=11.34). Notably, 80% (4/5) of the tetrad portfolios with the highest mortality impact included both cancer and COPD for male and female individuals. As cancer was consistently present in the triads and tetrads with the highest mortality impact, we separately examined the triads and tetrads among portfolios without cancer. The results are presented in Table 2. Upon excluding cancer, we observed that portfolios including dementia and schizophrenia were prominent in most of the triads and tetrads with the highest mortality impact. Among tetrads, the portfolios with the highest mortality impact for male individuals always involved osteoporosis paired with dementia or schizophrenia. In contrast, for female individuals, the tetrads with the highest mortality impact typically consisted of stroke in combination with dementia or schizophrenia.

**Table 1.** The 5 largest hazard ratios (HRs) for dyad, triad, and tetrad disease portfolios.

| Rank | Portfolio[a] | HR (99.9% CI)[b] | Individuals[c], n (%) |
|---|---|---|---|
| **Male individuals** | | | |
| **Dyads (n= 188,910 )** | | | |
| 1 | [CAN[d]] | 6.72 (6.06-7.45) | 6702 (3.55) |
| 2 | [DEM[e]] | 3.99 (3.59-4.43) | 1272 (0.67) |
| 3 | [SCH[f]] | 3.04 (2.85-3.24) | 888 (0.47) |
| 4 | [ST[g]] | 2.89 (2.66-3.14) | 5722 (3.03) |
| 5 | [COPD[h]] | 2.81 (2.55-3.10) | 7884 (4.17) |
| **Triads (n= 229,552 )** | | | |
| 1 | [CAN, SCH] | 13.26 (11.50-15.29) | 66 (0.03) |
| 2 | [CAN, COPD] | 11.34 (9.89-12.99) | 1356 (0.59) |
| 3 | [CAN, OP[i]] | 10.35 (9.01-11.90) | 433 (0.19) |
| 4 | [CAN, DEM] | 10.06 (8.38-12.07) | 131 (0.06) |
| 5 | [CAN, ST] | 9.87 (8.59-11.35) | 773 (0.34) |
| **Tetrads (n= 195,248 )** | | | |
| 1 | [CAN, COPD, SCH] | 19.21 (16.33-22.60) | 28 (0.01) |
| 2 | [CAN, SCH, ST] | 16.82 (14.14-20.01) | 14 (0.01) |
| 3 | [CAN, COPD, OP] | 16.40 (14.10-19.07) | 157 (0.08) |
| 4 | [CAN, COPD, ST] | 15.92 (13.29-19.07) | 168 (0.09) |
| 5 | [CAN, COPD, DEM] | 14.71 (11.59-18.67) | 30 (0.02) |
| **Female individuals** | | | |
| **D yads (n= 148,395 )** | | | |
| 1 | [CAN] | 7.59 (6.83-8.43) | 3559 (2.4) |
| 2 | [DEM] | 4.41 (3.98-4.89) | 1180 (0.8) |
| 3 | [ST] | 3.60 (3.27-3.97) | 3386 (2.28) |
| 4 | [COPD] | 3.57 (3.23-3.95) | 4335 (2.92) |
| 5 | [SCH] | 2.74 (2.56-2.92) | 663 (0.45) |
| **T riads (n= 190,272 )** | | | |
| 1 | [CAN, COPD] | 15.39 (13.51-17.53) | 622 (0.33) |
| 2 | [CAN, SCH] | 13.38 (11.70-15.31) | 58 (0.03) |
| 3 | [CAN, DEM] | 12.84 (10.60-15.56) | 90 (0.05) |
| 4 | [CAN, ST] | 12.65 (10.91-14.67) | 296 (0.16) |
| 5 | [CAN, DIA[j]] | 10.44 (9.24-11.80) | 251 (0.13) |
| **Tetrads (n= 177,755 )** | | | |
| 1 | [CAN, COPD, SCH] | 24.10 (20.45-28.41) | 14 (0.01) |
| 2 | [CAN, COPD, DEM] | 23.13 (17.89-29.91) | 13 (0.01) |
| 3 | [CAN, COPD, ST] | 22.80 (18.84-27.59) | 54 (0.03) |
| 4 | [CAN, DEM, ST] | 19.14 (15.03-24.37) | 20 (0.01) |
| 5 | [CAN, COPD, OP] | 17.57 (15.06-20.48) | 168 (0.09) |

[a]All portfolios contain the HD diagnosis.

[b]The reference group comprises male or female individuals with only heart disease (HD). HR estimates were aggregated on the log-hazard scale for male and female individuals across all educational attainment levels using weights corresponding to the number of individuals with each portfolio within

that subpopulation. Portfolios with <10 individuals were excluded.

[c]The number of unique male or female individuals who had exactly this combination of diseases at any time during the observation period. Percentages are among all male or female individuals observed with dyads, triads, and tetrads, respectively.

[d]CAN: cancer.

[e]DEM: dementia.

[f]SCH: schizophrenia.

[g]ST: stroke.

[h]COPD: chronic obstructive pulmonary disease.

[i]OP: osteoporosis.

[j]DIA: diabetes.

**Table 2.** The 5 largest hazard ratios (HRs) for dyad, triad, and tetrad disease portfolios excluding portfolios with cancer.

| Rank | Portfolio[a] | HR (99.9% CI)[b] | Number of individuals[c] |
|---|---|---|---|
| **Male individuals** | | | |
| **Dyads (n= 182,208 )** | | | |
| 1 | [DEM[d]] | 3.99 (3.59-4.43) | 1272 (0.7) |
| 2 | [SCH[e]] | 3.04 (2.85-2.24) | 888 (0.49) |
| 3 | [ST[f]] | 2.89 (2.66-3.14) | 5722 (3.14) |
| 4 | [COPD[g]] | 2.81 (2.55-3.10) | 7884 (4.33) |
| 5 | [OP[h]] | 2.47 (2.26-2.69) | 2341 (1.28) |
| **Triads (n= 206,638 )** | | | |
| 1 | [DEM, OP] | 8.58 (7.49-9.84) | 257 (0.12) |
| 2 | [DEM, ST] | 7.54 (6.58-8.65) | 380 (0.18) |
| 3 | [COPD, SCH] | 7.37 (6.58-8.24) | 177 (0.09) |
| 4 | [DEM, SCH] | 7.12 (6.34-8.00) | 228 (0.11) |
| 5 | [SCH, ST] | 6.50 (5.80-7.28) | 117 (0.06) |
| **Tetrads (n= 164,266 )** | | | |
| 1 | [DEM, OP, ST] | 13.37 (11.32-15.78) | 98 (0.06) |
| 2 | [DEM, OP, SCH] | 12.36 (10.46-14.61) | 52 (0.03) |
| 3 | [DEM, DIA[i], OP] | 12.09 (10.21-14.31) | 19 (0.01) |
| 4 | [COPD, DEM, OP] | 11.90 (10.00-14.16) | 42 (0.03) |
| 5 | [COPD, OP, SCH] | 11.72 (10.26-13.40) | 26 (0.02) |
| **Female individuals** | | | |
| **Dyads (n= 144,836 )** | | | |
| 1 | [DEM] | 4.41 (3.98-4.89) | 1180 (0.81) |
| 2 | [ST] | 3.60 (3.27-3.97) | 3386 (2.34) |
| 3 | [COPD] | 3.57 (3.23-3.95) | 4335 (2.99) |
| 4 | [SCH] | 2.74 (2.56-2.92) | 663 (0.46) |
| 5 | [DIA] | 2.31 (2.18-2.44) | 2939 (2.03) |
| **Triads (n= 174,861 )** | | | |
| 1 | [ST, DEM] | 9.77 (8.49-11.24) | 268 (0.15) |
| 2 | [COPD, DEM] | 8.68 (7.36-10.24) | 113 (0.06) |
| 3 | [COPD, SCH] | 8.44 (7.52-9.47) | 106 (0.06) |
| 4 | [ST, COPD] | 8.42 (7.27-9.75) | 324 (0.19) |
| 5 | [OP, DEM] | 7.96 (6.99-9.06) | 649 (0.37) |
| **Tetrads (n= 154,975 )** | | | |
| 1 | [COPD, DEM, ST] | 16.65 (13.54-20.47) | 32 (0.02) |
| 2 | [DEM, DIA, ST] | 15.04 (12.95-17.46) | 35 (0.02) |
| 3 | [COPD, SCH, ST] | 14.79 (12.56-17.41) | 12 (0.01) |
| 4 | [DEM, OP, ST] | 14.58 (12.36-17.20) | 142 (0.09) |
| 5 | [COPD, DEM, OP] | 13.86 (11.51-16.70) | 72 (0.05) |

[a]All portfolios contain the HD diagnosis.

[b]The reference group comprises male or female individuals with only heart disease (HD). HR estimates were aggregated on the log-hazard scale for male and female individuals across all educational attainment levels using weights corresponding to the number of individuals with each portfolio within

that subpopulation. Portfolios with <10 individuals were excluded.

[c]The number of unique male or female individuals who had exactly this combination of diseases at any time during the observation period. Percentages are among all male or female individuals observed with dyads, triads, and tetrads, respectively, excluding those with cancer.

[d]DEM: dementia.

[e]SCH: schizophrenia.

[f]ST: stroke.

[g]COPD: chronic obstructive pulmonary disease.

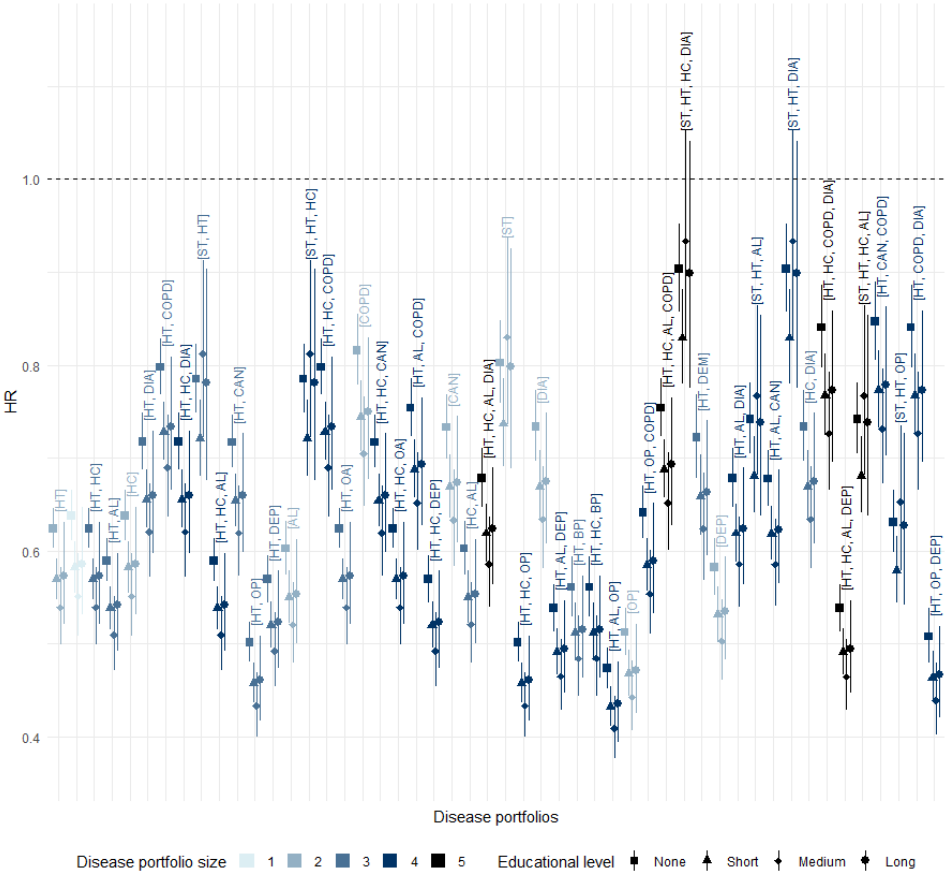
[h]OP: osteoporosis.

[i]DIA: diabetes.

### *Effect of Sex Across Socioeconomic Subpopulations*

The complex interactions at play indicate that the effect of sex on mortality varies by disease portfolio. This is illustrated in Figure 8, which presents HRs comparing female to male individuals across the 50 most prevalent disease portfolios at different educational levels. Overall, the figure shows a decrease in female mortality risk compared to male mortality risk, with most HRs falling below 1, ranging from 0.41 ([hypertension, allergies, osteoporosis]) to 0.93 ([stroke, high cholesterol, diabetes] and [stroke, hypertension, high cholesterol, diabetes]). However, the magnitude of this decrease varied across comorbidity patterns. For example, portfolios that included osteoporosis consistently showed HRs of <0.66, indicating a notably lower mortality risk for female individuals with these portfolios than for male individuals. Conversely, more complex disease portfolios that included stroke and diabetes—such as [stroke, hypertension, high cholesterol, diabetes] and [stroke, hypertension, diabetes]—had HRs closer to 1, suggesting only a slight reduction in female mortality hazard rate compared to male mortality hazard rate.

**Figure 8.** Hazard ratios (HRs) of female (vs male) sex by disease portfolio and educational attainment level. Estimates for the 50 most common disease portfolios are shown with 99.9% CIs. The estimates are presented for each of the educational attainment levels: none, short, medium, and long, indicated by different shapes and always in ascending order from none to long. The reference group comprises male individuals with the same disease portfolio and educational attainment level. The disease portfolios are ordered by prevalence from left to right, with [hypertension (HT)] being the most frequent disease portfolio. All portfolios contain the heart disease (HD) condition, so it is not labeled in the plot. Therefore, the disease portfolio without a label in the plot (the second from the left) corresponds to the disease portfolio with only HD. AL: allergies; BP: back pain; CAN: cancer; COPD: chronic obstructive pulmonary disease; DEM: dementia; DEP: depression; DIA: diabetes; HC: high cholesterol; OA: osteoarthritis; OP: osteoporosis; ST: stroke.



### *The Impact of COPD*

To illustrate that the effect of a single disease varies depending on the other diseases present in the portfolio, we estimated the effect of COPD in each observed disease portfolio in the population. The aggregated results are shown in Table 3 for male and female individuals of increasing disease portfolio size. The effect of COPD was greatest in triads (HR=2.81 for male individuals and 3.57 for female individuals) and generally higher in female than in male individuals. For increasing disease portfolio sizes, the aggregated effect of COPD decreased considerably with increasing disease portfolio sizes.

**Table 3.** Effect of chronic obstructive pulmonary disease (COPD) for increasing disease portfolio sizes. Each cell is the aggregated effect of COPD (ie, hazard ratio [HR] comparing the portfolio with and without COPD). The effects were aggregated on the log-hazard scale using weights determined based on the occurrence of each specific disease portfolio across the different educational attainment levels for each sex.

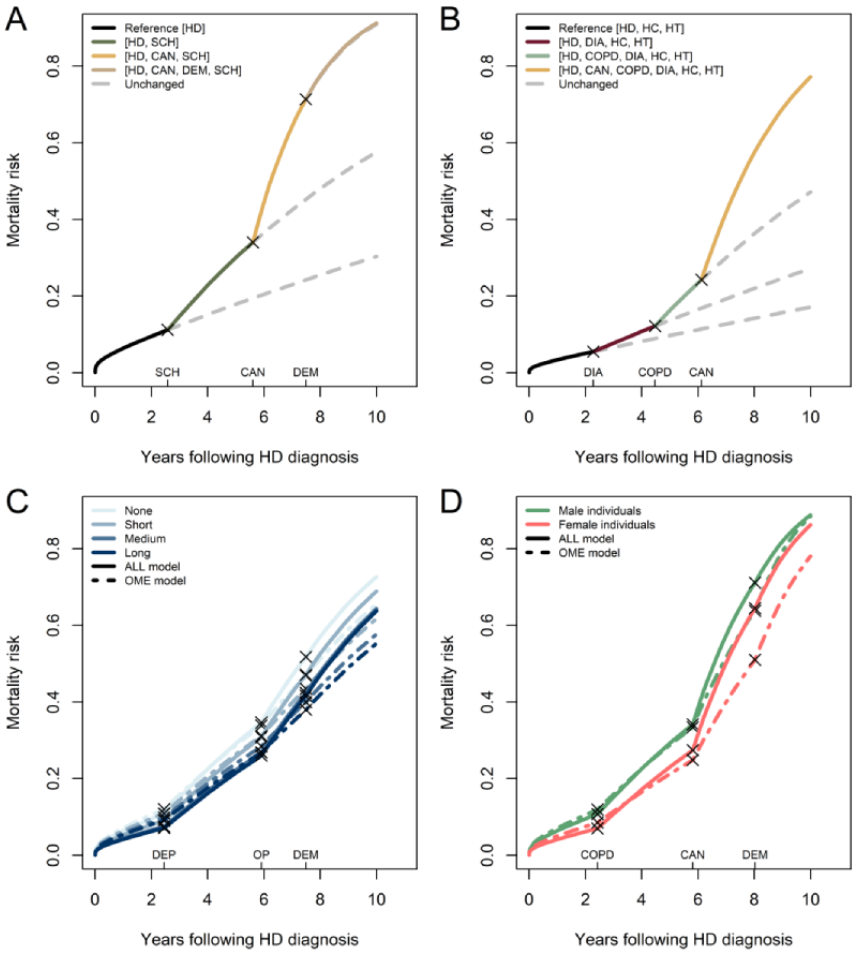
| Sex | Disease portfolio size | | | | | | |
|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| HR for male individuals | 2.81 | 2.98 | 2.74 | 2.50 | 2.27 | 2.08 | 1.91 |
| HR for female individuals | 3.57 | 3.77 | 3.43 | 3.08 | 2.75 | 2.45 | 2.19 |

## Scenarios

We present 4 scenarios in Figure 9 to illustrate how the ALL model's estimates translate to the risk scale. In Figure 9A, we show the first scenario, which consists of the trajectory of schizophrenia followed by cancer and then dementia. The figure illustrates an increase in the mortality rate with the additions of schizophrenia and cancer to the disease portfolio. However, when dementia diagnosis is obtained, its involvement in interactions prevents a substantial increase in the mortality rate compared to simply continuing undiagnosed. This is despite dementia being the disease with the second-highest mortality impact when considered in isolation (HR=3.99 for male individuals and 4.41 for female individuals; Table 1). The interaction effects between the diseases in the portfolio and dementia create a situation in which adding dementia does not further elevate the mortality hazard rate substantially.

Figure 9B shows a scenario that could resemble the disease trajectory of a male heavy smoker. In this scenario, the patient initially obtains HD diagnosis while also having hypertension and high cholesterol. Over the following years, the patient receives a diabetes diagnosis, which further elevates the mortality risk. The risk accelerates even more with the addition of a COPD diagnosis and, finally, a cancer diagnosis. In Figure 9C, a scenario showing the risk over time for a depression, osteoporosis, and dementia trajectory at different educational attainment levels for both the ALL and OME model is presented. A deviation between the ALL and OME models is most visible at the dementia disease, after which the risk in the ALL model accelerates compared to that in the OME model. In addition, the scenario visualizes that, despite the inverse social gradient of the disease portfolios on the log-hazard scale (Figure 7), lower educational attainment is still associated with a greater risk of death. Another scenario illustrating this relationship is presented in Figure 9D for a COPD, cancer, and dementia trajectory. In this scenario, we observe general increased mortality in male individuals compared to female individuals. However, due to the HRs of the disease portfolios being greater in female compared to male individuals (Table 1), the sex difference decreases over time.

**Figure 9.** Disease progression scenarios representing the mortality risk over time of a hypothetical (A) male individual with no education at mean age and calendar time who develops schizophrenia (SCH), cancer (CAN), and dementia (DEM) at 2.6, 5.6, and 7.5 years, respectively, following heart disease (HD) diagnosis; (B) male individual with no education who has hypertension (HT), high cholesterol (HC) at time of HD diagnosis and diabetes (DIA), chronic obstructive pulmonary disease (COPD), and CAN at 2.3, 4.8, and 6.1 years, respectively, following HD diagnosis; (C) male individual of varying educational attainment levels who develops depression (DEP), osteoporosis (OP), and DEM at 2.5, 5.9, and 7.5 years, respectively, following HD diagnosis under the all interactions (ALL) model (solid lines) and the additive only main effects (OME) model (dashed lines); and (D) male (green color) and female (red color) individual with no education who develops COPD, CAN, and DEM at 2.4, 5.8, and 8.0 years, respectively, following HD diagnosis under the ALL model (solid lines) and the additive OME model (dashed lines).




## Discussion

### Principal Findings

Patients with HD will often be diagnosed with other chronic diseases during their lifetime [2,5]. The effect of these co-occurring diseases on adverse outcomes is an important research focus as it is a clinically emerging challenge. In this study on the effect of disease portfolios on time until death, an extended Cox model allowing for time-varying covariates was applied to a large, longitudinal dataset encompassing all Danish adult patients with HD in the period from 1995 to 2015. We identified interactions through a model and data-driven variable selection procedure, revealing the severe diseases depression, stroke, COPD, dementia, and osteoporosis as involved in the most complex interactions. In addition, we estimated a simpler additive model consisting solely of main effects, which, on average, underestimated the effect of severe disease portfolios by a factor of 1.4. We did this to elucidate the importance of considering interaction effects when modeling the mortality risk associated with multiple chronic diseases. To the best of our knowledge, our work is the most extensive study examining the effect of co-occurring diseases on mortality among patients with HD.

We found that depression, stroke, COPD, dementia, and osteoporosis were involved in interaction relationships of the highest order, indicating that, when any of these diseases is added to the disease portfolio of the patient with HD, its risk contribution extensively depends on the other diseases already in the portfolio or the intrinsic variables describing the patient. These diseases were also identified under alternative variable selection procedures. Our comparisons between the interaction model and the simpler additive model showed differences in the magnitude of the effects for several disease portfolios. Overall, if interactions are not modeled, the average effect of disease portfolios on time until death appears underestimated for disease portfolios with >3 diseases (up to a factor of 1.4; Figure 4). For female individuals, this average underestimation also applied to disease portfolios of size 3. We observed an inverse socioeconomic gradient in the educational dimension for some of the most frequent and complex disease portfolios, where the greater the educational attainment level, the greater the associated HR of the disease portfolio (Figures 5-7;

Multimedia Appendix 8). We found that cancer was present in all cases in the disease portfolios with the highest mortality impact (Table 1). When considering disease portfolios with the highest mortality impact that did not include cancer, we observed that the psychiatric diseases schizophrenia and dementia frequently appeared in conjunction with osteoporosis for male individuals and in conjunction with stroke for female individuals (Table 2). Schizophrenia also often appeared with cancer among the disease portfolios with the highest mortality impact. These results highlight effect modification when multiple diseases co-occur in the patient with HD, and therefore, interventions should carefully evaluate the entire disease portfolio of the patient with HD.

### Effects and Interactions

The high complexity of the estimated interaction model is clearly illustrated in Figure 3. The figure shows the many dynamics between diseases at play in the HD population, where multiple chronic diseases are rampant. Depression, stroke, COPD, dementia, and osteoporosis were the chronic diseases included in the most complex interactions, also allowing for interactions between these and the patients' intrinsic factors. When considering interactions between chronic diseases exclusively (the DIO model), we observed that cancer and diabetes were also involved in the most complex interactions (Multimedia Appendix 3). Interactions with the intrinsic variables sex and age might trivially explain some of these interactions involving cancer and diabetes, which could be why they were not identified among the most complex interactions in the ALL model. Nevertheless, most interactions between individual chronic diseases identified in the ALL model variable selection were similarly discovered in either the stable or DIO model variable selections (Multimedia Appendix 5), indicating robustness in the detected interactions.

The consequences of modeling effects of interactions are meticulously visualized on the risk scale in the scenario illustrated in Figure 9A, where the addition of dementia does not change the risk profile of the hypothetical patient much as he already has the severe diseases schizophrenia and cancer along with HD. In fact, many of the effect modifications implied by the presence of interactions led to an attenuation of the combined effect of the diseases compared to their effects in an additive model. Biologically, this is reasonable as the considered patients are generally frail due to their HD, thereby causing the continued addition of chronic diseases to increase frailty before death eventually occurs. Our results showing the effect of COPD decreasing for increasing disease portfolio sizes support this finding (Table 3).

Our analysis showed that both the psychiatric diseases dementia and long-term depression were involved in the most complex interactions (Figure 3). Although not part of 5-way interactions, schizophrenia was involved in 4-way interactions with several other diseases. These high-order interaction effects in disease portfolios with psychiatric diseases complicate the interpretation of their impact on mortality as the effects of having these psychiatric diseases depend heavily on the other chronic diseases present in the portfolio, as well as on intrinsic factors such as age, sex, and socioeconomic position. From a biological point of view, this illustrates the interplay between somatic and psychiatric diseases concerning mortality [37,38]. Studies report increased prevalence and risk of psychiatric diagnoses for patients with cardiovascular diseases and their risk factors [39], and efforts should be made to improve these patients' psychological function. In addition, several studies indicate an increased mortality risk in psychiatric patients when comorbidities are present [7,37,38]. Indeed, we also found that the psychiatric diseases schizophrenia and dementia were present in the disease portfolios with the highest HRs (Tables 1 and 2). As a result, this study has substantial implications for the priority of identifying psychiatric manifestations of multimorbidity among patients with HD as mortality risk is heavily modified when these diagnoses are present, at least among the chronic diseases and the population considered in this study.

Cancer was present in all portfolio dyads, triads, and tetrads with the highest HRs (Table 1). This finding is supported by previous studies reporting that most deaths from cardiovascular disease occur in patients diagnosed with breast, bladder, and prostate cancer [40]. However, the cancer diagnosis in our study encapsulated a larger spectrum of cancer conditions. Among the triads and tetrads with the highest mortality impact, cancer was often present with schizophrenia. However, when considering portfolios excluding cancer, dyads with dementia had a higher mortality impact. Previous research shows higher cancer mortality rates in individuals with schizophrenia, often attributed to factors such as nonadherence to treatment, diagnostic overshadowing, and limited collaboration between medicine and psychiatry [41]. For patients with HD, our results highlight these combinations of diseases as having some of the most substantial mortality impacts.

We note that, among the variables identified in higher-order interactions, Figure 7 and Multimedia Appendix 8 show differences in effects when comparing estimates from models with and without interactions. These contrasts emphasize the importance of considering the complete disease portfolio of a patient with HD when assessing risk. Our findings show that, when interactions are not recognized, the model underestimates the effect of severe diseases such as cancer, stroke, and COPD while overestimating the effect of less severe diseases such as high cholesterol and allergies (Figure 5). A previous study demonstrated the adverse impact of ignoring statistical interactions in epidemiologic studies, showing a potential bias in main effect parameter estimates [33], which could be a reason for these observed differences. As the underestimation of effects asserted itself even for disease portfolios of small size, it could be attributed to the first few manifestations of multimorbidity (ie, the first diseases developed after HD) being more important for survival than later. While the risk continuously increases with the addition of diagnoses, the individual disease effects do not combine additively. As a result, some patients might reach a high risk profile with just a few diagnoses, trivializing the extra effect of obtaining a new diagnosis, as illustrated by the scenario in Figure 9A. The situation illustrated in Figure 9A with the mortality risk not changing with the addition of a (on its own) deadly chronic disease can only be modeled when interactions are allowed. We speculate that the simple additive model breaks down due to situations such as these,

compensating the underestimation of the effect of severe diseases with an overestimation of the effect of more common, less severe diseases. While it was observed that, on average, the additive model underestimated the effect of disease portfolios (Figure 4), it is essential to mention that the individual disease portfolio effect differences were aggregated across the HD population.

In this study, we observed an apparent negative effect of the high cholesterol diagnosis, indicating increased survival relative to an individual without the disease. This artifact can be attributed to the phenomenon that some individuals diagnosed with HD who are also diagnosed with high cholesterol are likely being treated with lipid-modifying agents such as statins, which have many beneficial properties such as cholesterol reduction and anti-inflammatory effects [42,43]. Despite having an additional diagnosis, these individuals diagnosed with HD might represent a less frail part of the HD population who might have a higher degree of health literacy, thus being more aware of their conditions and receiving attention from their general practitioners. Another possible explanation is our use of diagnosis time instead of the time of actual disease onset, which was unknown. High cholesterol is a condition in which a considerable amount of time may pass before diagnosis [44], and among those patients with HD who are undiagnosed, some may have the disease but not be undergoing treatment. It is also essential to consider other consequences of multimorbidity. Increased survival relative to an individual without a particular disease may appear beneficial at first glance. However, it is crucial to recognize that an additional chronic disease introduces new challenges, such as new medication management, consultations with general practitioners and specialists, and potential functional impairments. It is essential to remember that increased survival in these cases does not necessarily equate to improved quality of life.

We found a more pronounced effect in disease portfolios including osteoporosis in male individuals compared to female individuals (Figures 5, 6, and 8; Table 1). Notably, despite the generally higher prevalence of osteoporosis in female individuals compared to male individuals, it is well documented that male individuals diagnosed with osteoporosis experience higher mortality rates than their female counterparts [45]. Our study reaffirms this observation within a nationwide HD population.

Our findings revealed an inverse socioeconomic gradient for some disease portfolios, where the isolated effect of disease portfolios generally increased as educational attainment levels rose (Figures 5-7; Multimedia Appendix 8). Thus, the higher educated the patient, the higher the mortality hazard rate of the disease portfolio compared to a person of the same educational level with only HD. It is widely known that individuals with higher levels of education enjoy better overall health and lower mortality hazard rates than people with lower levels of education [46]. Consequently, given that the reference patient with HD who was not multimorbid was generally healthier in the subpopulation with the highest educational attainment, it is plausible that those who do become multimorbid in this subpopulation experience a comparatively higher relative mortality hazard rate. Hence, when interpreting this inverse social gradient, it is important to bear in mind that the HR reflects the increased relative mortality hazard rate associated with having a specific multimorbid disease portfolio compared to only having HD. Importantly, the inverse social gradient does not directly translate to increased mortality with higher educational level on the risk scale, as illustrated in Figure 7C. Social disparities are extensively documented across various aspects of multimorbidity, including prevalence [21], health care use [47], and transitions between disease portfolios [28]. Our results contribute to this by revealing an inverse social gradient concerning the isolated effect of combinations of chronic diseases on mortality within a nationwide HD population.

As clinical practice, such as guidelines, screening, testing, and treatment for chronic diseases, evolved over the period from 1995 to 2015, our analysis was adjusted for calendar time at HD diagnosis. We systematically assessed the influence of calendar time on the most frequently observed disease portfolios. Generally, we observed increased survival for patients diagnosed more recently compared to earlier (of the 100 most common portfolios, n=98, 98%). However, an inverse trend indicating decreased survival over calendar time was observed for a few disease portfolios, particularly for the portfolio [dementia] and, in many cases, when dementia was combined with diabetes or stroke. It is well known that demographic changes have caused an increase in the prevalence of dementia over the years [48], but as the model is conditional on the disease portfolio, an increased prevalence of dementia over time does not in itself explain the result. We currently lack an explanation for this result and plan to further investigate it in future research.

## Interpretations

This study illustrates that the complexity of addressing the effects of multiple chronic conditions in a large, temporal dataset requires consideration of the individual's complete disease portfolio. The extended Cox model used throughout this work was chosen because it allows for modeling time-varying variables in a survival context. In addition, it has the advantage of making no assumptions regarding the distribution of the survival times (ie, the underlying hazard function is left unspecified [49]). However, a few assumptions were made about the hazard function, namely, the relationship between covariates and the hazard function. By examining Schoenfeld residuals, we found that, in some cases, the proportional hazard assumption was not fully supported [31], meaning that the effects might vary across time. Therefore, it is essential to interpret the presented effects as weighted averages of the true, possibly time-varying effects across the entire observation period [50]. There are previous studies on the effect of multimorbidity on time to death within HD populations [5,7]. However, the analyses conducted in these studies do not acknowledge that a patient's multimorbidity state is likely to change dynamically through time (ie, that it is time dependent). The differences in prevalence at time $t$=0 and the end of the observation period (Figure 2) in this study illustrate much progression in disease portfolios. Thus, it is essential to consider this when conducting a temporal statistical analysis. When interpreting effects, it is crucial to keep the population in mind. As the study population was selected and followed up on from the time of HD diagnosis, the individuals considered were generally ill compared to, for

example, an individual without any chronic diseases. Furthermore, with Denmark being a European welfare state, the population differs from those of many other countries where individuals may have to pay for examinations; thus, the effects might not be directly comparable due to variations in treatment accessibility.

It is crucial to elaborate further on the contrasts associated with the presented effect estimates. The estimates presented compare a patient with HD who is not multimorbid to a patient with HD who is multimorbid with a specific disease portfolio. In the OME model, the effect of comparing, for example, a patient with HD diagnosed with cancer and COPD to a patient with HD who is not multimorbid would be the same as comparing a patient with HD who also has cancer, COPD, and depression to a patient with HD who also has depression. In other words, the effect of a disease combination in an additive model can be interpreted as having the specific combination of diagnoses in the disease portfolio versus not having it. However, in the presence of high-order interactions, the interpretation is only the increased (or decreased) effect comparing an individual with the particular disease portfolio to an individual without it. This is due to the possibility of interactions with other variables, which modify the effects of the disease combination.

The scenarios in Figure 9 were created to illustrate the workings of the extended Cox model by illustrating how the model estimates the mortality risk over time for the hypothetical individuals diagnosed with HD. However, one should be careful in interpreting these scenarios. They cannot be used prognostically to forecast as time points of portfolio expansions are never known at the time of HD diagnosis as that would be conditioning on future events. These scenarios were solely constructed to represent how the model depicts the mortality risk of a “typical” patient with HD over time. The figures help illustrate how the interaction effects on the log-hazard scale relate to the risk of mortality on the probability scale.

For the results presented in this paper, it is essential to emphasize that the effects and interactions uncovered represent associations, not causal relationships. While our results provide valuable insights into the relationships among the chronic diseases, they should be interpreted as observational associations, which can be informative for hypothesis generation and risk assessment for individual portfolios. Furthermore, a considerable group of individuals had missing educational attainment information in this study. In our analyses, we modeled missing values as separate categories. We also estimated the final ALL model under the multiple imputation framework [51], which led to similar results as those presented.

### Strengths and Weaknesses

The main strength of this study is the entire Danish population of individuals diagnosed with HD observed over a long period using register data. Danish register data are generally of high quality and fully representative of the entire Danish population [52]. In addition, the use of algorithmic diagnoses processing both *International Classification of Diseases, 10th Revision*, diagnosis history and Anatomical Therapeutic Chemical medicine history ensured that the HD population covered both the primary and secondary parts of the Danish health care system. However, there are several limitations associated with this study. Given the observational nature of this study, our results do not enable us to draw causal conclusions. In addition, despite the algorithmic diagnoses previously being shown to be reliable [18], a chronic disease’s true onset comes before diagnosis. This is less of a challenge when diagnoses are considered in a cross-sectional study than in a longitudinal setting. Therefore, as time stamps for true disease onsets are not possible, it is crucial to interpret the longitudinal effects associated with a diagnosis in the context of exactly a diagnosis (ie, the detection of the disease), where the individual may have been ill for some time before that.

## Conclusions

In conclusion, we emphasize the importance of considering a patient’s entire disease portfolio when assessing or modeling risk, avoiding oversimplified silo-based generalizations about the effect of individual diseases. This study highlights the importance of modeling interaction effects when chronic diseases co-occur. Omitting these interactions can result in underestimation of the elevated mortality risk associated with multimorbidity in patients with HD. Through our analysis of a comprehensive nationwide longitudinal dataset of 766,596 patients with HD, we identified sex-related and socioeconomic disparities in disease portfolio HRs. Notably, an inverse socioeconomic gradient was systematically observed for the most common and complex disease portfolios, meaning an increased mortality hazard rate with multimorbidity relative to no multimorbidity as educational attainment level increases. However, absolute mortality risk still decreased with increasing educational attainment due to baseline effects of education. Cancer was present in all disease portfolios with the highest mortality impact. Excluding cancer, disease portfolios including psychiatric chronic diseases were of the highest mortality impact. We identified interactions among all considered co-occurring chronic diseases. We found that stroke, osteoporosis, COPD, dementia, and depression were integral components of the most complex interactions of the highest order. When these chronic diseases co-occur in the patient with HD, their contribution to the patient’s risk profile depends on multiple factors, encouraging a holistic view of the patient’s entire disease portfolio along with their demographic and socioeconomic risk factors.

## Acknowledgments

This study was funded by Greater Copenhagen Health Science Partners as part of the Clinical Academic Group Prognostication of Acute Recovery Capacity – in an Aging Population. The funder played no role in the study design; data collection, analysis,

and interpretation; or writing of this manuscript. The authors would like to thank Guillermina Eslava for stimulating discussions and useful suggestions on earlier drafts of the manuscript.

## Data Availability

The data used in this study are not publicly available as they consist of sensitive, individual-level information in the form of national register data. According to the Danish data protection legislation, the authors are not allowed to share these sensitive data upon request. Instead, the data are available for research purposes upon request to the Danish Health Authority.

## Authors' Contributions

AF, AS, and NNH developed the design and concept of this study. HGJL, NNH, and OA made substantial contributions to the preparation of data. NNH made all software implementations. NNH conducted the statistical analysis with assistance from AS. NNH wrote the initial draft of the manuscript in collaboration with AF and AS. All authors made substantial contributions to the interpretation of the results. All authors read and approved the final version of the manuscript.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Algorithmic diagnoses. Algorithms used to define the 15 diagnoses.
[DOCX File , 17 KB - cardio_v9i1e57749_app1.docx ]

Multimedia Appendix 2
Prevalence of diagnoses according to sex.
[DOCX File , 17 KB - cardio_v9i1e57749_app2.docx ]

Multimedia Appendix 3
Graphical representation of disease-disease interactions in the disease interactions only model. A ribbon connects chronic diseases that have any significant interaction ($P$<.001) between them. The connection’s width corresponds to the number of individuals diagnosed with HD developing both diseases throughout the observation period. The ribbon’s color represents the highest-order interaction relationship between 2 diseases. The ribbon chart is ordered by number of connections between diseases, starting from allergies with 5 connections all the way to cancer, which interacts with all the additional diseases
[PNG File , 397 KB - cardio_v9i1e57749_app3.png ]

Multimedia Appendix 4
Graphical representation of disease-disease interactions in the stable model. A ribbon connects chronic diseases that have any significant interaction ($P$<.001) between them. The connection’s width corresponds to the number of individuals diagnosed with HD developing both diseases throughout the observation period. The ribbon’s color represents the highest-order interaction relationship between 2 diseases. The ribbon chart is ordered by number of connections between diseases, starting from high cholesterol with 1 connection all the way to chronic obstructive pulmonary disease, which interacts with 11 of the additional diseases. Back pain does not interact with any chronic disease in this model.
[PNG File , 297 KB - cardio_v9i1e57749_app4.png ]

Multimedia Appendix 5
Diagnosis-diagnosis interactions identified across the all interactions model, the disease interactions only model, and the stable model. A cell in the table indicates under which models arising from the different variable selection procedures an interaction between the row and column condition is identified. Due to symmetry, only half of the table is presented.
[DOCX File , 20 KB - cardio_v9i1e57749_app5.docx ]

Multimedia Appendix 6
Male hazard ratios (HRs) for the 10 most common disease portfolio dyads, triads, tetrads, and pentads. The results are presented for the all interactions model at the 4 educational attainment levels (none, short, medium, and long) and correspond to the situation presented in Figure 5. The reference group comprises male individuals with only heart disease and the corresponding educational attainment level. Results are also presented for the additive only main effects model. In each disease portfolio group, the disease portfolio HR estimates are presented in order of prevalence, with the upper rows being more prevalent than the lower rows.
[DOCX File , 23 KB - cardio_v9i1e57749_app6.docx ]

Multimedia Appendix 7
Female hazard ratios (HRs) for the 10 most common disease portfolio dyads, triads, tetrads, and pentads. The results are presented for the all interactions model at the 4 educational attainment levels (none, short, medium, and long) and correspond to the situation presented in Figure 6. The reference group comprises female individuals with only heart disease and the corresponding educational attainment level. Results are also presented for the additive only main effects model. In each disease portfolio group, the disease portfolio HR estimates are presented in order of prevalence, with the upper rows being more prevalent than the lower rows.
[DOCX File , 23 KB - cardio_v9i1e57749_app7.docx ]

Multimedia Appendix 8
Effects of disease portfolio dyads (A), triads (B), tetrads (C), and pentads (D) involving stroke, osteoporosis, chronic obstructive pulmonary disease, dementia, and depression. Effects are shown for female individuals of varying educational attainment levels at the log-hazard rate scale. Comparisons are made to a female individual of the corresponding educational attainment level who only has heart disease (HD). Effects are presented for the all interactions model (different shades of blue) and the only main effects model (red). All comparisons are made at mean age and calendar time. The HD condition is present in all disease portfolios.
[PNG File , 168 KB - cardio_v9i1e57749_app8.png ]

## References


1. Nowbar AN, Gitto M, Howard JP, Francis DP, Al-Lamee R. Mortality from ischemic heart disease. Circ Cardiovasc Qual Outcomes 2019 Jun;12(6):e005375 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.118.005375] [Medline: 31163980]
2. Forman DE, Maurer MS, Boyd C, Brindis R, Salive ME, Horne FM, et al. Multimorbidity in older adults with cardiovascular disease. J Am Coll Cardiol 2018 May 15;71(19):2149-2161 [FREE Full text] [doi: 10.1016/j.jacc.2018.03.022] [Medline: 29747836]
3. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for understanding health and health services. Ann Fam Med 2009;7(4):357-363 [FREE Full text] [doi: 10.1370/afm.983] [Medline: 19597174]
4. Buddeke J, Bots ML, van Dis I, Visseren FL, Hollander M, Schellevis FG, et al. Comorbidity in patients with cardiovascular disease in primary care: a cohort study with routine healthcare data. Br J Gen Pract 2019 Jun;69(683):e398-e406 [FREE Full text] [doi: 10.3399/bjgp19X702725] [Medline: 31064742]
5. Gimeno-Miguel A, Gracia Gutiérrez A, Poblador-Plou B, Coscollar-Santaliestra C, Pérez-Calvo JI, Divo MJ, et al. Multimorbidity patterns in patients with heart failure: an observational Spanish study based on electronic health records. BMJ Open 2019 Dec 23;9(12):e033174 [FREE Full text] [doi: 10.1136/bmjopen-2019-033174] [Medline: 31874886]
6. Chen YH, Karimi M, Rutten-van Mölken MP. The disease burden of multimorbidity and its interaction with educational level. PLoS One 2020;15(12):e0243275 [FREE Full text] [doi: 10.1371/journal.pone.0243275] [Medline: 33270760]
7. Sariaslan A, Sharpe M, Larsson H, Wolf A, Lichtenstein P, Fazel S. Psychiatric comorbidity and risk of premature mortality and suicide among those with chronic respiratory diseases, cardiovascular diseases, and diabetes in Sweden: a nationwide matched cohort study of over 1 million patients and their unaffected siblings. PLoS Med 2022 Jan 27;19(1):e1003864 [FREE Full text] [doi: 10.1371/journal.pmed.1003864] [Medline: 35085232]
8. van Deursen VM, Urso R, Laroche C, Damman K, Dahlström U, Tavazzi L, et al. Co-morbidities in patients with heart failure: an analysis of the European Heart Failure Pilot Survey. Eur J Heart Fail 2014 Jan;16(1):103-111 [FREE Full text] [doi: 10.1002/ejhf.30] [Medline: 24453099]
9. Eriksson B, Wändell P, Dahlström U, Näsman P, Lund LH, Edner M. Comorbidities, risk factors and outcomes in patients with heart failure and an ejection fraction of more than or equal to 40% in primary care- and hospital care-based outpatient clinics. Scand J Prim Health Care 2018 Jun;36(2):207-215 [FREE Full text] [doi: 10.1080/02813432.2018.1459654] [Medline: 29633886]
10. Haue AD, Armenteros JJ, Holm PC, Eriksson R, Moseley PL, Køber LV, et al. Temporal patterns of multi-morbidity in 570157 ischemic heart disease patients: a nationwide cohort study. Cardiovasc Diabetol 2022 May 31;21(1):87 [FREE Full text] [doi: 10.1186/s12933-022-01527-3] [Medline: 35641964]
11. Willadsen TG, Siersma V, Nicolaisdóttir DR, Køster-Rasmussen R, Reventlow S, Rozing M. The effect of disease onset chronology on mortality among patients with multimorbidity: a Danish nationwide register study. J Multimorb Comorb 2022 Aug 21;12:26335565221122025 [FREE Full text] [doi: 10.1177/26335565221122025] [Medline: 36032184]
12. France EF, Wyke S, Gunn JM, Mair FS, McLean G, Mercer SW. Multimorbidity in primary care: a systematic review of prospective cohort studies. Br J Gen Pract 2012 Apr;62(597):e297-e307 [FREE Full text] [doi: 10.3399/bjgp12X636146] [Medline: 22520918]
13. Multimorbidity: a priority for global health research. The Academy of Medical Sciences. 2018 Apr. URL: https://acmedsci.ac.uk/file-download/82222577 [accessed 2025-04-04]
14. Muth C, Kirchner H, van den Akker M, Scherer M, Glasziou PP. Current guidelines poorly address multimorbidity: pilot of the interaction matrix method. J Clin Epidemiol 2014 Nov;67(11):1242-1250. [doi: 10.1016/j.jclinepi.2014.07.004] [Medline: 25216898]

15. Muth C, Blom JW, Smith SM, Johnell K, Gonzalez-Gonzalez AI, Nguyen TS, et al. Evidence supporting the best clinical management of patients with multimorbidity and polypharmacy: a systematic guideline review and expert consensus. J Intern Med 2019 Mar;285(3):272-288 [FREE Full text] [doi: 10.1111/joim.12842] [Medline: 30357955]
16. Pedersen CB. The Danish civil registration system. Scand J Public Health 2011 Jul;39(7 Suppl):22-25. [doi: 10.1177/1403494810387965] [Medline: 21775345]
17. Data protection rules as a trust-enabler in the EU and beyond – taking stock. European Commission. 2019 Jul 24. URL: https://www.europeansources.info/record/data-protection-rules-as-a-trust-enabler-in-the-eu-and-beyond-taking-stock/ [accessed 2025-04-04]
18. Robinson KM, Lau CJ, Jeppesen M, Vind AB, Glu¨mer C. Kroniske sygdomme - hvordan opgøres kroniske sygdomme? Psykiatrien Region Sjælland.: `; 2011. URL: https://research.regionh.dk/en/publications/kroniske-sygdomme-hvordan-opg%C3%B8res-kroniske-sygdomme [accessed 2025-04-04]
19. Christensen AI, Davidsen M, Ekholm O, Pedersen PV, Juel K. Danskernes sundhed - den nationale sundhedsprofil 2013. Sundhedsstyrelsen. 2014 Mar 5. URL: https://www.sst.dk/da/udgivelser/2014/Danskernes-Sundhed---Den-Nationale-Sundhedsprofil-2013 [accessed 2025-04-04]
20. Jensen HA, Davidsen M, Ekholm O, Christensen AI. Danskernes sundhed - den nationale sundhedsprofil 2017. Sundhedsstyrelsen. 2018 Mar. URL: https://www.sst.dk/da/udgivelser/2018/Danskernes-Sundhed-Den-Nationale-Sundhedsprofil-2017 [accessed 2025-04-04]
21. Schiøtz ML, Stockmarr A, Høst D, Glümer C, Frølich A. Social disparities in the prevalence of multimorbidity - a register-based population study. BMC Public Health 2017 May 10;17(1):422 [FREE Full text] [doi: 10.1186/s12889-017-4314-8] [Medline: 28486983]
22. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol 2015 Nov 17;7:449-490 [FREE Full text] [doi: 10.2147/CLEP.S91125] [Medline: 26604824]
23. Mors O, Perto GP, Mortensen PB. The Danish psychiatric central research register. Scand J Public Health 2011 Jul;39(7 Suppl):54-57. [doi: 10.1177/1403494810395825] [Medline: 21775352]
24. Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. Scand J Public Health 2011 Jul;39(7 Suppl):38-41. [doi: 10.1177/1403494810394717] [Medline: 21775349]
25. Andersen JS, Olivarius ND, Krasnik A. The Danish National Health Service Register. Scand J Public Health 2011 Jul;39(7 Suppl):34-37. [doi: 10.1177/1403494810394718] [Medline: 21775348]
26. Jensen VM, Rasmussen AW. Danish education registers. Scand J Public Health 2011 Jul;39(7 Suppl):91-94. [doi: 10.1177/1403494810394715] [Medline: 21775362]
27. Helweg-Larsen K. The Danish register of causes of death. Scand J Public Health 2011 Jul;39(7 Suppl):26-29. [doi: 10.1177/1403494811399958] [Medline: 21775346]
28. Holm NN, Frølich A, Andersen O, Juul-Larsen HG, Stockmarr A. Longitudinal models for the progression of disease portfolios in a nationwide chronic heart disease population. PLoS One 2023 Apr 20;18(4):e0284496 [FREE Full text] [doi: 10.1371/journal.pone.0284496] [Medline: 37079591]
29. Berrington de González A, Cox DR. Interpretation of interaction: a review. Ann Appl Stat 2007 Dec 1;1(2):371-385. [doi: 10.1214/07-aoas124]
30. Cox DR. Regression models and life-tables. J R Stat Soc Ser B Methodol 1972;34(2):187-202. [doi: 10.1111/j.2517-6161.1972.tb00899.x]
31. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994;81(3):515-526. [doi: 10.1093/biomet/81.3.515]
32. Brambor T, Clark W, Golder M. Understanding interaction models: improving empirical analyses. Polit Anal 2017 Jan 04;14(1):63-82. [doi: 10.1093/pan/mpi014]
33. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. The effect of ignoring statistical interactions in regression analyses conducted in epidemiologic studies: an example with survival analysis using Cox proportional hazards regression model. Epidemiology (Sunnyvale) 2015 Feb;6(1):216 [FREE Full text] [doi: 10.4172/2161-1165.1000216] [Medline: 27347436]
34. Meinshausen N, Bühlmann P. Stability selection. J R Stat Soc Series B Stat Methodol 2010;72(4):417-473. [doi: 10.1111/j.1467-9868.2010.00740.x]
35. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Methodol 2018 Dec 05;58(1):267-288. [doi: 10.1111/j.2517-6161.1996.tb02080.x]
36. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med 1997 Feb 28;16(4):385-395. [doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3]
37. Koyanagi A, Köhler-Forsberg O, Benros ME, Munk Laursen T, Haro JM, Nordentoft M, et al. Mortality in unipolar depression preceding and following chronic somatic diseases. Acta Psychiatr Scand 2018 Dec;138(6):500-508. [doi: 10.1111/acps.12899] [Medline: 29761489]
38. Kugathasan P, Stubbs B, Aagaard J, Jensen SE, Munk Laursen T, Nielsen RE. Increased mortality from somatic multimorbidity in patients with schizophrenia: a Danish nationwide cohort study. Acta Psychiatr Scand 2019 Oct;140(4):340-348. [doi: 10.1111/acps.13076] [Medline: 31355419]

39. Zhang Y, Chen Y, Ma L. Depression and cardiovascular disease in elderly: current understanding. J Clin Neurosci 2018 Jan;47:1-5. [doi: 10.1016/j.jocn.2017.09.022] [Medline: 29066229]
40. Sturgeon KM, Deng L, Bluethmann SM, Zhou S, Trifiletti DM, Jiang C, et al. A population-based study of cardiovascular disease mortality risk in US cancer patients. Eur Heart J 2019 Dec 21;40(48):3889-3897 [FREE Full text] [doi: 10.1093/eurheartj/ehz766] [Medline: 31761945]
41. González-Rodríguez A, Labad J, Seeman MV. Schizophrenia and cancer. Curr Opin Support Palliat Care 2020 Sep;14(3):232-238. [doi: 10.1097/SPC.0000000000000512] [Medline: 32701859]
42. Niazi M, Galehdar N, Jamshidi M, Mohammadi R, Moayyedkazemi A. A review of the role of statins in heart failure treatment. Curr Clin Pharmacol 2020;15(1):30-37 [FREE Full text] [doi: 10.2174/1574884714666190802125627] [Medline: 31376825]
43. Lee MM, Sattar N, McMurray JJ, Packard CJ. Statins in the prevention and treatment of heart failure: a review of the evidence. Curr Atheroscler Rep 2019 Jul 27;21(10):41 [FREE Full text] [doi: 10.1007/s11883-019-0800-z] [Medline: 31350612]
44. EAS Familial Hypercholesterolaemia Studies Collaboration (FHSC). Global perspective of familial hypercholesterolaemia: a cross-sectional study from the EAS Familial Hypercholesterolaemia Studies Collaboration (FHSC). Lancet 2021 Nov 06;398(10312):1713-1725. [doi: 10.1016/S0140-6736(21)01122-3] [Medline: 34506743]
45. Rinonapoli G, Ruggiero C, Meccariello L, Bisaccia M, Ceccarini P, Caraffa A. Osteoporosis in men: a review of an underestimated bone condition. Int J Mol Sci 2021 Feb 20;22(4):2105 [FREE Full text] [doi: 10.3390/ijms22042105] [Medline: 33672656]
46. Davies NM, Dickson M, Davey Smith G, van den Berg GJ, Windmeijer F. The causal effects of education on health outcomes in the UK Biobank. Nat Hum Behav 2018 Feb;2(2):117-125 [FREE Full text] [doi: 10.1038/s41562-017-0279-y] [Medline: 30406209]
47. Frølich A, Ghith N, Schiøtz M, Jacobsen R, Stockmarr A. Multimorbidity, healthcare utilization and socioeconomic status: a register-based study in Denmark. PLoS One 2019;14(8):e0214183 [FREE Full text] [doi: 10.1371/journal.pone.0214183] [Medline: 31369580]
48. Nichols E, Szoeke CEI, Vollset SE, Abbasi N, Abd-Allah F, Abdela J, et al. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol Elsevier 2019;18(1):88-106 [FREE Full text]
49. Andersen PK, Borgan Ø, Gill RD, Keiding N. Statistical Models Based on Counting Processes. New York, NY: Springer; 1993.
50. Stensrud MJ, Hernán MA. Why test for proportional hazards? JAMA 2020 Apr 14;323(14):1401-1402. [doi: 10.1001/jama.2020.1267] [Medline: 32167523]
51. van Buuren S. Flexible Imputation of Missing Data. New York, NY: Chapman & Hall; 2012.
52. Erlangsen A, Fedyszyn I. Danish nationwide registers for public health and health-related research. Scand J Public Health 2015 Jun;43(4):333-339. [doi: 10.1177/1403494815575193] [Medline: 25759376]


## Abbreviations

**COPD:** chronic obstructive pulmonary disease
**DIO:** disease interactions only
**HD:** heart disease
**HR:** hazard ratio
**ICD-10:** International Statistical Classification of Diseases, Tenth Revision
**LASSO:** least absolute shrinkage and selection operator
**OME:** only main effects


*Edited by A Coristine; submitted 26.02.24; peer-reviewed by E Amini-Salehi, Y Zhu; comments to author 20.01.25; revised version received 10.03.25; accepted 10.03.25; published 25.04.25.*

*Please cite as:*
*Holm NN, Frølich A, Dominguez H, Dalhoff KP, Juul-Larsen HG, Andersen O, Stockmarr A*
*Co-Occurring Diseases and Mortality in Patients With Chronic Heart Disease, Modeling Their Dynamically Expanding Disease Portfolios: Nationwide Register Study*
*JMIR Cardio 2025;9:e57749*
*URL: https://cardio.jmir.org/2025/1/e57749*
*doi:10.2196/57749*
*PMID:*

Original Paper

# Lipid Profile and Apolipoprotein B Serum Levels in the Vietnamese Population With Newly Diagnosed Elevated Low-Density Lipoprotein Cholesterol and Association With the Single-Nucleotide Variant rs676210: Cross-Sectional Study

Quyen Thuy Nguyen[1*], MD; An Viet Tran[1], MD, PhD; Bao The Nguyen[2], MSc, MD; Hoa Thai Nguyen[1], MD, PhD; Nhung Thi Hong Thai[1], MD, PhD; Hen Huu Phan[3*], MD, PhD

[1]Faculty of Medicine, Can Tho University of Medicine and Pharmacy, Can Tho, Vietnam

[2]University of Medicine and Pharmacy, Hue University, Thua Thien Hue, Vietnam

[3]Cho Ray Hospital, Ho Chi Minh, Vietnam

[*]these authors contributed equally

**Corresponding Author:**
An Viet Tran, MD, PhD
Faculty of Medicine
Can Tho University of Medicine and Pharmacy
179 Nguyen Van Cu, An Khanh Ward, Ninh Kieu District
Can Tho, 902510
Vietnam
Phone: 84 939060818
Email: tvan@ctump.edu.vn

## Abstract

**Background:** Apolipoprotein B (APOB) rs676210 polymorphism has been associated with altered lipid metabolism and cardiovascular risk in various populations; however, data from Vietnamese populations remain limited.

**Objective:** This study aimed to investigate the association of the APOB rs676210 variant with lipid profiles among Vietnamese individuals newly diagnosed with elevated low-density lipoprotein cholesterol (LDL-C).

**Methods:** A cross-sectional study was conducted among 69 Vietnamese adults newly diagnosed with elevated LDL-C (≥130 mg/dL) at a tertiary hospital in Southern Vietnam. Participants were genotyped for APOB rs676210 using real-time polymerase chain reaction (PCR) with allele-specific probes. Lipid profile components, including LDL-C, high-density lipoprotein cholesterol (HDL-C), non–HDL-C, and ApoB, were compared across genotype groups (AA vs GA/GG) and alleles (A vs G). Statistical analyses involved t tests, chi-square tests, and multivariable linear regression adjusted for age, sex, the BMI, and diabetes. $P<.05$ was considered statistically significant.

**Results:** Of the 69 participants, 32 (46.4%) carried the AA genotype, while 37 (53.6%) carried the GA or the GG genotype. The AA genotype was associated with significantly higher LDL-C (mean 5.19, SD 0.95, vs mean 4.37, SD 0.97, mmol/L; $P<.001$), non–HDL-C (mean 5.94, SD 1.08, vs mean 5.31, SD 1.22 mmol/L; $P=.03$), and ApoB (mean 149.5, SD 26.3, vs mean 136.9, SD 15.2, mg/dL; $P=.02$) and lower HDL-C (mean 1.26, SD 0.31, vs mean 1.44, SD 0.39, mmol/L; $P=.03$) compared to the GA/GG genotype. Allele-based analysis showed that carriers of the A allele (98/138, 71%) also had higher LDL-C (mean 4.91, SD 1.02, vs mean 4.36, SD 0.97, mmol/L; $P=.004$) and ApoB (mean 145.6, SD 23.2, vs mean 135.9, SD 16.0, mg/dL; $P=.02$) than G allele carriers (40/138, 29%). These associations remained significant after multivariate adjustment.

**Conclusions:** APOB rs676210 polymorphism is associated with significant differences in lipid profiles among Vietnamese adults with elevated LDL-C. Specifically, the A allele and the AA genotype confer a more atherogenic profile, suggesting potential utility as a genetic marker in lipid screening and personalized cardiovascular risk management in this population.

***(JMIR Cardio 2025;9:e76850)*** doi:10.2196/76850

**KEYWORDS**

APOB rs676210 polymorphism; APOB; ApoB; apolipoprotein B; lipid profile; elevated LDL-C; low-density lipoprotein cholesterol

## Introduction

Cardiovascular disease (CVD) remains a leading cause of mortality worldwide, and dyslipidemia is a key modifiable risk factor contributing to this burden. Elevated low-density lipoprotein cholesterol (LDL-C), in particular, is a significant causal factor in atherosclerotic cardiovascular disease (ASCVD) [1]. In many Asian populations, including Vietnam, the prevalence of lipid disorders is high, with roughly one-third to nearly half of adults meeting the criteria for elevated LDL-C levels [2]. In this context, there is a growing body of research in Vietnam focusing on genetic factors related to CVD, highlighting the emergence of genomics as a relevant field [3-5]. Understanding genetic contributions to lipid abnormalities in Vietnamese patients is thus increasingly important to improve cardiovascular risk stratification and guide personalized interventions.

Apolipoprotein B (ApoB) is the primary protein component of LDL and other atherogenic lipoproteins, and it plays a crucial role in the assembly, transport, and cellular uptake of these particles. ApoB is a major structural protein of very low-density lipoproteins (VLDLs) and low-density lipoproteins (LDLs), mediating their interaction with cellular receptors [6]. Given its central function, genetic variations in the *APOB* gene can significantly impact lipid metabolism. Indeed, numerous single-nucleotide polymorphisms (SNPs) in *APOB* have been associated with altered plasma lipid levels and increased atherosclerosis risk [7]. One such polymorphism is rs676210, a c.8216G>A variant in exon 26 of *APOB* that results in a proline-to-leucine substitution at codon 2739 (p.Pro2739Leu) [6]. This missense mutation is of particular interest biologically, as it is expected to induce a functional change in the ApoB-100 protein structure. The variant has been reported to influence LDL particle characteristics; for example, rs676210 has been linked to the susceptibility of LDL to oxidative modification [8]. Oxidized LDL has a pathogenic role in plaque formation; thus, such a genetic effect could directly affect ASCVD risk.

Epidemiologically, prior evidence suggests that rs676210 may be relevant to interindividual differences in lipid profiles and coronary risk. Genome-wide analyses have identified rs676210 as a locus associated with plasma lipoprotein traits [8,9]. Notably, the minor (A) allele of rs676210 was associated with a more favorable lipid profile (lower total cholesterol [TC], triglycerides [TGs], and LDL-C and higher high-density lipoprotein cholesterol [HDL-C]) in a large cohort study [8]. In addition, this SNP has been implicated in cardiovascular outcomes: for instance, it was associated with myocardial infarction risk in Chinese populations, likely mediated by hyperlipidemia and higher ApoB levels [10,11]. However, findings across studies have not been consistent, and data on rs676210 are scarce in Southeast Asian groups, such as the Vietnamese.

In Vietnam and other middle-income countries, resources for comprehensive genotyping (eg, whole-genome sequencing) are limited [12]. Therefore, focusing on common functional SNPs, such as rs676210, is a practical approach to investigate genetic predisposition to dyslipidemia in these populations. Given the rising burden of hypercholesterolemia in Vietnam [13], it is pertinent to investigate whether genetic polymorphisms, such as *APOB* rs676210, contribute to interindividual variations in LDL-C levels and related lipid indices among the Vietnamese population. In contrast, statin therapy has been shown to be effective in improving lipid profiles in high-risk patients [14,15]. However, the response to statins can vary considerably among individuals, possibly due to underlying genetic factors [6,10]. In this context, the identification of lipid-related polymorphisms may provide a useful foundation for tailoring preventive and therapeutic interventions. This study thus aimed to explore the association of rs676210 with plasma lipid parameters (LDL-C, non-HDL-C, HDL-C, and ApoB levels) in a cohort of Vietnamese patients newly diagnosed with elevated LDL-C. By doing so, we sought to clarify the biological and clinical significance of this variant in an Asian middle-income country, where identifying key genetic markers could aid in improving the screening and management of dyslipidemia.

## Methods

### Study Design and Population

This cross-sectional descriptive study was conducted among adults undergoing routine health checkups at Can Tho University of Medicine and Pharmacy Hospital, a major medical center in Can Tho City, the economic, cultural, and health care hub of the Mekong Delta region in Southern Vietnam. The study was implemented from April 2023 to February 2025. Participants were identified during annual occupational health examinations and were newly diagnosed with elevated LDL-C. None had received lipid-lowering therapy prior to enrollment.

A nonprobability convenience sampling method was used. Eligible participants were individuals aged 18 years or older who were diagnosed with elevated LDL-C, defined as LDL-C≥130 mg/dL (3.4 mmol/L), based on the previous literature and partly because the National Cholesterol Education Program (Adult Treatment Panel III), or NCEP ATP III, guidelines have identified this threshold as predictive of an increased risk of ASCVD [16-18]. Strict exclusion criteria were applied to ensure participant safety and homogeneity of the study population. Patients who were currently taking medications known to affect blood lipid levels, such as corticosteroids, immunosuppressants, oral contraceptives, and CYP3A4 inhibitors (including diltiazem, rifamycins, cyclosporine, erythromycin, itraconazole, ketoconazole, HIV protease inhibitors, fosamprenavir, and ritonavir), were excluded from the study. Additionally, individuals with a history of secondary dyslipidemia-inducing conditions, such as nephrotic syndrome and hypothyroidism, as well as those with familial hypercholesterolemia, were also excluded from the study.

A post hoc power analysis was conducted based on the observed LDL-C difference of 0.88 mmol/L ($\beta$=.877, SE 0.237) and a pooled SD of 0.96 mmol/L. At a significance level of $\alpha$=.05, the achieved statistical power ($1 - \beta$) was estimated at 84%.

### Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics

Committee of Biomedical Research at Can Tho University of Medicine and Pharmacy (approval no: 23.052.HV-ĐHYDCT). Prior to enrollment, written informed consent was obtained from all participants after they had been clearly informed of the study's objectives, procedures, potential risks, and confidentiality safeguards. All data collected were anonymized using unique identifier codes to ensure participant privacy, and no personally identifiable information was included in the dataset or manuscript. Participants received no financial or material compensation for their involvement, as the study was conducted during routine health screenings. Furthermore, no individual-identifiable images or data were presented in the manuscript or supplementary materials; therefore, image consent forms were not required.

## Data Collection

A standardized questionnaire was used to collect clinical and lifestyle data, including age, sex, history of smoking, alcohol abuse, sedentary lifestyle, diabetes mellitus, and hypertension. Anthropometric measurements, including height, weight, and the BMI, were directly measured using standardized procedures. Height was measured to the nearest 0.1 cm, while weight was recorded to the nearest 0.1 kg. The BMI was subsequently calculated as weight in kilograms divided by the square of height in meters ($kg/m^2$). The classification of overweight and obesity was based on the National Institutes of Health and World Health Organization guidelines for the Asian population [19]. Hypertension was defined as previously diagnosed or newly diagnosed according to the 2023 European Society of Cardiology guidelines [20]. Blood pressure measurement procedures were standardized using the 2020 International Society of Hypertension guidelines [21]. Serum urea was quantified using the glutamate dehydrogenase (GLDH) method, which used the enzyme GLDH to measure the reduction of reduced nicotinamide adenine dinucleotide (NADH), a process directly proportional to the urea concentration in the sample. Serum creatinine was measured using the Jaffé kinetic method (the only method available in Vietnam), where creatinine reacts with an alkaline picrate reagent to form a yellow-orange complex [22]. The rate of complex formation is proportional to the creatinine level when compared with a standard. This was performed on the Abbott Architect c4000 automated biochemistry analyzer using the Biolabo reagents from Abbott. Fasting blood glucose and hemoglobin A1c (HbA1c) were tested using the Abbott Architect c4000 automated biochemistry analyzer. Diabetes mellitus was defined as previously diagnosed or newly diagnosed based on the 2023 criteria of the American Diabetes Association or the current use of glucose-lowering medications [23]. After an overnight fast lasting 12 hours, venous blood samples were collected in the morning and processed according to standard laboratory procedures. The serum was separated by centrifugation and used for the determination of lipid profile components, including TC, TGs, HDL-C, and LDL-C. Among these, TC, TGs, and HDL-C were directly measured using enzymatic colorimetric methods on an automated clinical chemistry analyzer, with reagent kits provided with Biolabo reagents. The LDL-C concentration was calculated indirectly using Friedewald's formula in cases where TG levels were below 400 mg/dL: LDL-C = TC – HDL-C – (TGs/5) [24]. In our study, all participants had TG levels <400 mg/dL, satisfying the prerequisite for the formula's validity. In addition, serum ApoB levels were also quantified using immunoturbidimetric or chemiluminescent immunoassay methods, depending on the available analytical platform, using Biolabo reagents.

## Genotyping of APOB rs676210 Polymorphism

### *Genomic DNA Extraction*

Genomic DNA was extracted from peripheral blood using the Toppure Blood DNA Extraction Kit (ABT), following the manufacturer's protocol. Briefly, 200 µL of whole blood was mixed with 400 µL of BL buffer and 20 µL of proteinase K, followed by incubation at 72 °C for 10 minutes. After ethanol precipitation, the lysate was transferred to a silica spin column, washed sequentially with Wash Buffer (WB)1 and WB2, and eluted in 50 µL of EB buffer. The purified DNA was stored at –20 °C until use.

### *Real-Time PCR Genotyping*

Genotyping of the *APOB* rs676210 polymorphism was performed using a custom-designed allele-specific TaqMan assay. Each 25 µL reaction contained 2.5 µL of 10× polymerase chain reaction (PCR) buffer, 1 µL of primer mix (forward and reverse, 10 µM each), 1 µL of either a 6-carboxyfluorescein (FAM)- or a hexachloro-fluorescein (HEX)-labeled probe (5 µM), 1.5 µL of genomic DNA (10 ng/µL), and nuclease-free water to adjust the volume. The reaction used the EZ PCR Mix (Phu Sa Genomics), a ready-to-use premix containing Taq DNA polymerase, deoxynucleoside triphosphates (dNTPs), and $MgCl_2$ (final $Mg^2$ concentration=2.0 mM).

Amplification was performed on a CFX Opus 96 Real-Time PCR system (Bio-Rad Laboratories) under the following cycling conditions: initial denaturation at 95 °C for 5 minutes, followed by 35 cycles of denaturation at 95 °C for 25 seconds and annealing/extension at 60 °C for 45 seconds (fluorescence acquisition), and a final extension at 72 °C for 5 minutes. Fluorescence signals were captured and analyzed using CFX Maestro v2.3 software (Bio-Rad Laboratories).

### *Primers and Probes*

Primers and allele-specific dual-labeled hydrolysis probes were adapted from Abdulfattah and Al-Awadi [6], with minor sequence modifications to enhance allele specificity. The oligonucleotide sequences were as follows:

- Forward primer: 5′-TGTGTGTGAGATGTGGGGAA-3′
- Reverse primer: 5′-GGGATCTGAAGGTGGAGGAC-3′
- FAM-labeled probe (G allele): 5′-FAM-TCTGGTATGTGAAGGTCAGGA-3′-BHQ1
- HEX-labeled probe (A allele): 5′-HEX-TTCTGATATGTGAAGGTCAGGAAC-3′-BHQ1

All oligonucleotides were synthesized by Macrogen Inc. Primers were desalted, and probes were purified using high-performance liquid chromatography (HPLC).

### *Genotype Interpretation*

Allelic discrimination was based on fluorescence threshold detection:

- Homozygous G/G (Pro/Pro): Only the FAM signal exceeded the threshold.
- Homozygous A/A (Leu/Leu): Only the HEX signal exceeded the threshold.
- Heterozygous G/A (Pro/Leu): Both FAM and HEX signals were detected.

No internal control gene was used in the reaction. Duplicate reactions for consistency confirmed genotype calls, and Sanger sequencing was performed on 15% of the total samples to validate the accuracy of genotyping results.

### Data Analysis

Data were checked for completeness and accuracy before analysis. All variables were complete; no missing data were observed. Data collected in this study were encoded and processed using R version 4.3.3 (R Foundation for Statistical Computing), using libraries such as *tidyverse*, *dplyr*, *ggplot2*, *table1*, *compareGroups*, and *pROC* [25]. Data were checked for completeness and accuracy before analysis, with missing or invalid cases excluded. Outliers and abnormal values were detected using frequency plots and basic statistical checks via functions such as *filter()*, *drop_na()*, and *replace_na()*. Categorical variables were numerically coded for statistical processing using the *mutate()* function from *dplyr*, and at least 10% of the data underwent cross-checking using the *sample_n()* function to detect entry errors.

Categorical variables were summarized as frequencies (counts and percentages) using the *table1* package. For continuous variables, data distribution was assessed using the Shapiro-Wilk test, with the *shapiro.test()* function, and values were presented as means (SDs). Prior to statistical analysis, the Kolmogorov-Smirnov and Shapiro-Wilk tests, along with graphical distribution assessments (eg, histograms and quintile-quintile [Q-Q] plots), were used to determine data distribution, guiding the selection of appropriate statistical methods (parametric or nonparametric). When continuous data followed a normal distribution, comparisons between 2 groups were conducted using the Student *t* test, with the *t.test()* function, when the variance assumption was met. Additionally, the Levene test was performed to assess the homogeneity of variances. All continuous lipid variables were confirmed to satisfy the assumptions of normality and homogeneity of variances, validating the use of the Student *t* test for group comparisons. For categorical variables, group comparisons were performed using the chi-square test, with the *chisq.test()* function, when the expected frequency assumptions were satisfied; otherwise, the Fisher exact test, with the *fisher.test()* function, was applied for 2×2 contingency tables, and the Fisher-Freeman-Halton exact test was used for larger contingency tables.

For visualizing continuous data distributions, violin plots were generated using the *ggplot2* package, which provides a clear graphical representation of the data distribution, highlighting the variations in LDL-C and ApoB levels across different genotype groups and allele groups.

The associations between *APOB* rs676210 polymorphism (genotype and allele frequencies) and lipid profile parameters (LDL-C, HDL-C, non–HDL-C, and ApoB) were evaluated using statistical tests. Results are presented in statistical tables, and all tests were 2-tailed, with a significance level of $P<.05$.

## Results

### Participant Details

Figure 1 illustrates the flow diagram of participant recruitment and selection. A total of 69 patients with newly diagnosed elevated LDL-C levels were enrolled in the study. Regarding general characteristics, most participants were female, with a female-to-male ratio of approximately 1.38. Most patients were middle-aged, with a mean age of 54.54 (SD 11.64) years, and approximately three-quarters (n=50, 72.5%) of the cohort were overweight or obese. In terms of lifestyle behaviors, about one-quarter (n=16, 23.2%) of participants were current smokers. Concerning medical history, roughly one-third (n=21, 30.4%) had a diagnosis of hypertension, while one-sixth (n=10, 14.5%) had concomitant diabetes mellitus. A relatively high proportion of patients (n=29, 42%) reported a sedentary lifestyle. Additionally, one-sixth (n=12, 17.4%) reported alcohol abuse. With regard to genotype distribution, 32 (46.3%) participants had the AA genotype, 34 (49.3%) had the GA genotype, and 3 (4.4%) had the GG genotype. There were no statistically significant differences in baseline characteristics between the AA and the combined GA+GG genotype groups ($P<.05$), as shown in Table 1.

**Figure 1.** Flow diagram of study participant selection. LDL-C: low-density lipoprotein cholesterol; CYP3A4: cytochrome P450 3A4.

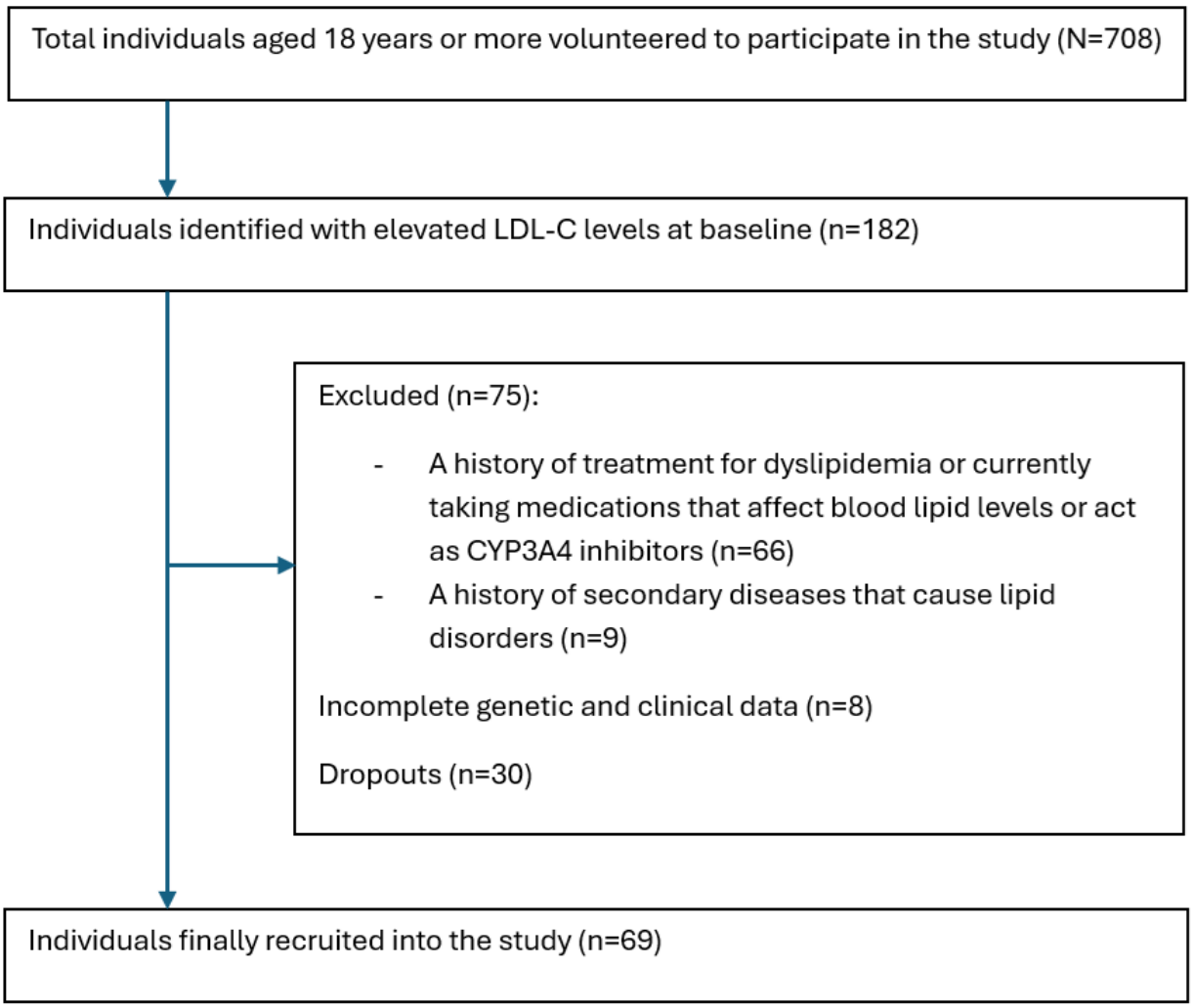


**Table 1.** Baseline characteristics of study participants stratified by genotype.

| Characteristics | Genotype | | | *P* value |
|---|---|---|---|---|
| | AA (n=32) | GA+GG (n=37) | Total (N=69) | |
| **Age (years)** | | | | |
| <60, n (%) | 23 (71.9) | 25 (67.6) | 48 (69.6) | .69[a] |
| ≥60, n (%) | 9 (28.1) | 12 (32.4) | 21 (30.4) | —[b] |
| Mean (SD) | 53.28 (12.61) | 55.62 (10.79) | 54.54 (11.64) | .41[c] |
| **Sex, n (%)** | | | | |
| Male | 14 (43.8) | 15 (40.5) | 29 (42) | .79[a] |
| Female | 18 (56.3) | 22 (59.5) | 40 (58) | — |
| Weight (kg), mean (SD) | 62.94 (6.06) | 60.73 (8.76) | 61.75 (7.66) | .24[c] |
| Height (cm), mean (SD) | 160 (7.13) | 159.24 (8.42) | 159.59 (7.8) | .69[c] |
| BMI (kg/m$^2$), mean (SD) | 24.61 (2.09) | 23.89 (2.43) | 24.23 (2.29) | .19[c] |
| **Overweight and obesity, n (%)** | | | | |
| Yes | 25 (78.1) | 25 (67.6) | 50 (72.5) | .33[a] |
| No | 7 (21.9) | 12 (32.4) | 19 (27.5) | — |
| **Smoking, n (%)** | | | | |
| Yes | 9 (28.1) | 7 (18.9) | 16 (23.2) | .37[a] |
| No | 23 (71.9) | 30 (81.1) | 53 (76.8) | — |
| **Alcohol abuse, n (%)** | | | | |
| Yes | 6 (18.8) | 6 (16.2) | 12 (17.4) | .78[a] |
| No | 26 (81.3) | 31 (83.8) | 57 (82.6) | — |
| **Sedentary lifestyle, n (%)** | | | | |
| Yes | 14 (43.8) | 15 (40.5) | 29 (42.0) | .79[a] |
| No | 18 (56.3) | 22 (59.5) | 40 (58.0) | — |
| **Diabetes mellitus, n (%)** | | | | |
| Yes | 4 (12.5) | 6 (16.2) | 10 (14.5) | .74[d] |
| No | 28 (87.5) | 31 (83.8) | 59 (85.5) | — |
| **Hypertension, n (%)** | | | | |
| Yes | 9 (28.1) | 12 (32.4) | 21 (30.4) | .70[a] |
| No | 23 (71.9) | 25 (67.6) | 48 (69.6) | — |

[a]Comparison of the differences are given according to the Pearson chi-square test (statistical significance at $P<.05$).

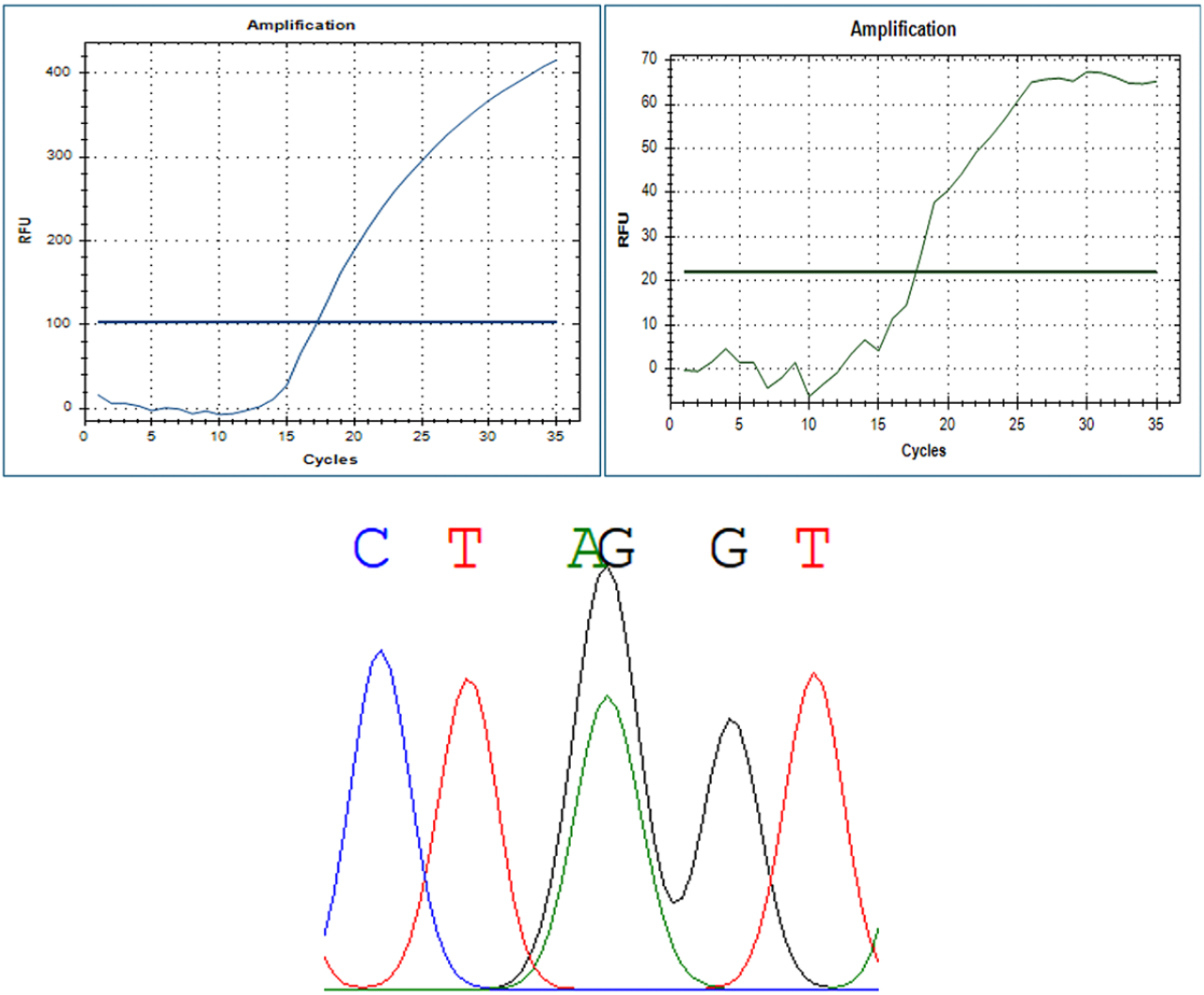
[b]Not applicable.

[c]Comparison of the differences are given according to the independent samples test (statistical significance at $P<.05$).

[d]Comparison of the differences are given according to the Fisher exact test (statistical significance at $P<.05$).

Genotyping quality control showed a 100% call rate, with all 10 blinded duplicate samples and 15% of Sanger-validated samples demonstrating 100% concordance. The genotype distribution of the *APOB* rs676210 polymorphism did not deviate significantly from Hardy-Weinberg equilibrium ($P=.15$), supporting assay reliability and population representativeness. Figure 2 shows the concordance between real-time PCR and Sanger sequencing results.

**Figure 2.** Concordant genotyping results using Sanger sequencing and real-time PCR (patient 24, GA genotype). PCR: polymerase chain reaction; RFU: relative fluorescence units.




Analysis of the lipid profile among the 69 study participants revealed that individuals with the AA genotype tend to have lower HDL-C levels compared to those with GA and GG genotypes (mean 1.26, SD 0.31, vs mean 1.44, SD 0.39, mmol/L; $P$=.03). Additionally, concentrations of LDL-C, non–HDL-C, and ApoB were significantly higher in the AA genotype group compared to the GA and GG groups (all $P$<.05), as shown in Table 2.

**Table 2.** Lipid profile, ApoBa concentrations, and other indices stratified by genotype.

| Parameters | Genotype | | | *P* value |
|---|---|---|---|---|
| | AA (n=32) | GA+GG (n=37) | Total (N=69) | |
| **TC[b] (mmol/L)** | | | | |
| Normal, n (%) | 0 | 4 (10.8) | 4 (5.8) | .21[c] |
| Borderline high, n (%) | 6 (18.8) | 6 (16.2) | 12 (17.4) | —[d] |
| High, n (%) | 26 (81.3) | 27 (73) | 53 (76.8) | — |
| Mean (SD) | 7.2 (1.18) | 6.75 (1.25) | 6.96 (1.23) | .14[e] |
| **TGs[f] (mmol/L)** | | | | |
| Normal, n (%) | 10 (31.3) | 13 (35.1) | 23 (33.3) | .89[c] |
| Borderline high, n (%) | 22 (68.8) | 23 (62.2) | 45 (65.2) | — |
| High, n (%) | 0 (0) | 1 (2.7) | 1 (1.4) | — |
| Mean (SD) | 2.44 (0.67) | 2.5 (0.82) | 2.47 (0.75) | .73[e] |
| **HDL-C[g] (mmol/L)** | | | | |
| Normal, n (%) | 26 (81.3) | 35 (94.6) | 61 (88.4) | .13[c] |
| Decreased, n (%) | 6 (18.8) | 2 (5.4) | 8 (11.6) | — |
| Mean (SD) | 1.26 (0.31) | 1.44 (0.39) | 1.36 (0.37) | .03[e] |
| **LDL-C[h] (mmol/L)** | | | | |
| Borderline high, n (%) | 4 (12.5) | 16 (43.2) | 20 (29.0) | .005[c] |
| High, n (%) | 28 (87.5) | 21 (56.8) | 49 (71.0) | — |
| Mean (SD) | 5.19 (0.95) | 4.37 (0.97) | 4.75 (1.04 ) | <.001[e] |
| Non–HDL-C (mmol/L), mean (SD) | 5.94 (1.08) | 5.31 (1.22) | 5.6 (1.19) | .03[e] |
| ApoB (mg/dL), mean (SD) | 149.5 (26.3) | 136.92 (15.21) | 142.75 (21.86) | .02[e] |
| Hemoglobin (g/dL), mean (SD) | 13.69 (1.26) | 13.86 (1.42) | 13.78 (1.34) | .59[e] |
| HbA1c[i] (%), mean (SD) | 6.05 (1.82) | 6.25 (1.75) | 6.16 (1.77) | .65[e] |
| Glucose (mmol/L), mean (SD) | 5.84 (1.84) | 5.94 (1.85) | 5.9 (1.83) | .83[e] |
| Creatinine (µmol/L), mean (SD) | 72.3 (16.01) | 70.32 (15.48) | 71.24 (15.65) | .60[e] |
| Urea (mmol/L), mean (SD) | 5.2 (1.71) | 5.1 (1.57) | 5.15 (1.63) | .80[e] |

[a]ApoB: apolipoprotein B.

[b]TC: total cholesterol.

[c]Comparison of the differences are given according to the Fisher-Freeman-Halton exact test (statistical significance at $P<.05$).

[d]Not applicable.

[e]Comparison of the differences are given according to the independent samples test (statistical significance at $P<.05$).

[f]TG: triglyceride.

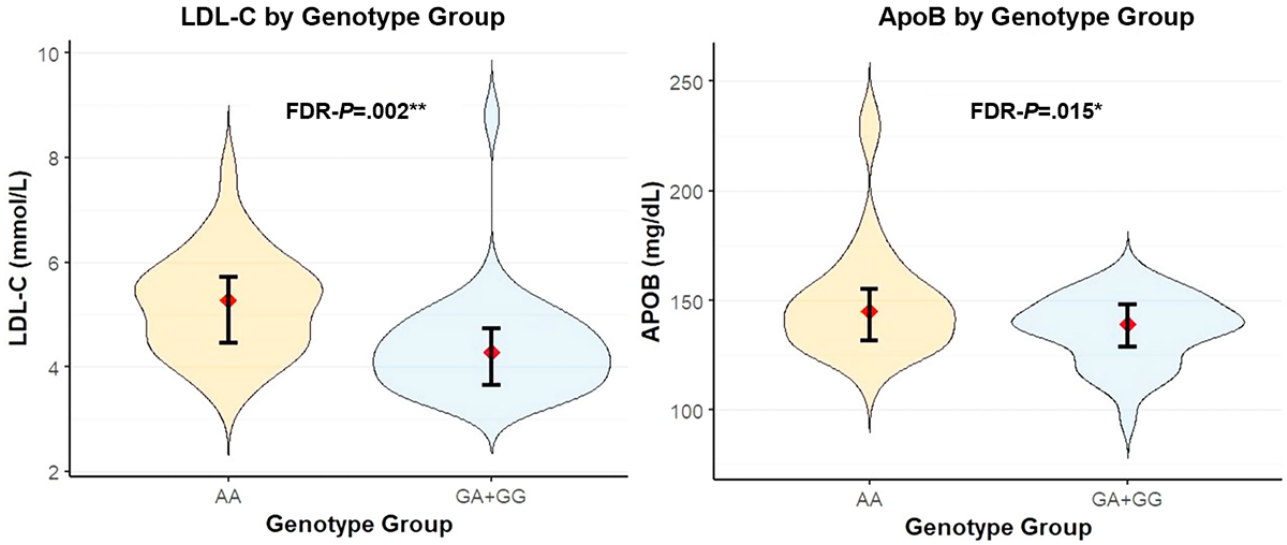
[g]HDL-C: high-density lipoprotein cholesterol.

[h]LDL-C: low-density lipoprotein cholesterol.

[i]HbA1c: hemoglobin A1c.

Figure 3 illustrates the distribution of LDL-C (mmol/L) and ApoB (mg/dL) levels by genotype. Individuals with the AA genotype exhibited higher LDL-C and ApoB concentrations than the GA+GG group. This suggests a potential association between the AA genotype and elevated atherogenic lipid parameters.

**Figure 3.** Violin plot of LDL-C and ApoB concentrations by genotypes group. ApoB: apolipoprotein B; FDR-P: false discovery rate–adjusted *P* value; LDL-C: low-density lipoprotein cholesterol.




To account for potential confounders, multivariable linear regression was performed for lipid parameters by genotype, adjusting for age, sex, the BMI, and diabetes mellitus. As shown in Table 3, associations between the AA genotype and increased LDL-C, non–HDL-C, and ApoB concentrations remained statistically significant after adjustment, while the inverse relationship with HDL-C was attenuated.

**Table 3.** Association between the APOB[a] rs676210 genotype and lipid profile outcomes: crude and adjusted estimates with multivariable linear regression.

| Outcome | Crude estimate *P* value[b] | Model 1[c] adjusted β (SE; 95% CI) | Model 1 *P* value | Model 2[d] adjusted β (SE; 95% CI) | Model 2 *P* value | FDR-*P*[e] |
|---|---|---|---|---|---|---|
| HDL-C[f] | .03 | –0.157 (0.087; –0.330 to 0.016) | .07 | –0.135 (0.088; –0.310 to 0.040) | .13 | .07 |
| LDL-C[g] | <.001 | 0.877 (0.237; 0.403 to 1.352) | <.001 | 0.922 (0.242; 0.439 to 1.405 | <.001 | .002 |
| Non–HDL-C | .03 | 0.673 (0.288; 0.098 to 1.249) | .02 | 0.745 (0.292; 0.162 to 1.328) | .01 | .04 |
| ApoB | .02 | 11.406 (5.163; 1.092 to 21.720) | .03 | 12.490 (5.247; 2.004 to 22.975) | .02 | .04 |

[a]*APOB*: apolipoprotein B.

[b]Crude *P* values from unadjusted comparisons between genotype groups.

[c]Model 1: age, sex, and BMI.

[d]Model 2: age, sex, BMI, and diabetes.

[e]FDR-P: false discovery rate–adjusted *P* value calculated for model 1 comparisons to account for multiple testing.

[f]HDL-C: high-density lipoprotein cholesterol.

[g]LDL-C: low-density lipoprotein cholesterol.

Each of the 69 participants contributed 2 alleles, resulting in a total of 138 alleles analyzed for allele-based comparisons. Stratified by genotype, there were no statistically significant differences in baseline characteristics, including age, sex, overweight-obesity status, smoking, alcohol abuse, sedentary lifestyle, diabetes mellitus, and hypertension between allele groups (all *P*>.05), as shown in Table 4.

Allele-based analysis indicated that carriers of the A allele are associated with lower levels of HDL-C and higher levels of LDL-C, non–HDL-C, and ApoB compared to carriers of the G allele (all *P*<.05), as shown in Table 5.

**Table 4.** Baseline characteristics of 69 participants stratified by allele.

| Characteristics | Allele | | | *P* value |
|---|---|---|---|---|
| | A (n=98) | G (n=40) | Total (N=138) | |
| **Age group (years)** | | | | |
| <60, n (%) | 70 (71.4) | 26 (65) | 96 (69.6) | .46[a] |
| ≥60, n (%) | 28 (28.6) | 14 (35) | 42 (30.4) | —[b] |
| Mean (SD) | 53.87 (11.96) | 56.18 (10.63) | 54.54 (11.6) | .29[c] |
| **Sex, n (%)** | | | | |
| Male | 43 (43.9) | 15 (37.5) | 58 (42) | .49[a] |
| Female | 55 (56.1) | 25 (62.5) | 80 (58) | — |
| Weight (kg), mean (SD) | 62.6 (6.7) | 59.68 (9.3) | 61.75 (7.63) | .08[c] |
| Height (cm), mean (SD) | 160.02 (7.34) | 158.55 (8.75) | 159.59 (7.77) | .30[c] |
| BMI (kg/m$^2$), mean (SD) | 24.46 (2.15) | 23.66 (2.54) | 24.23 (2.29) | .06[c] |
| **Overweight and obesity, n (%)** | | | | |
| Yes | 74 (75.5) | 26 (65) | 100 (72.5) | .21[a] |
| No | 24 (24.5) | 14 (35) | 38 (27.5) | — |
| **Smoking, n (%)** | | | | |
| Yes | 25 (25.5) | 7 (17.5) | 32 (23.2) | .31[a] |
| No | 73 (74.5) | 33 (82.5) | 106 (76.8) | — |
| **Alcohol abuse, n (%)** | | | | |
| Yes | 18 (18.4) | 6 (15.0) | 24 (17.4) | .64[a] |
| No | 80 (81.6) | 34 (85.0) | 114 (82.6) | — |
| **Sedentary lifestyle, n (%)** | | | | |
| Yes | 42 (42.9) | 16 (40.0) | 58 (42.0) | .76[a] |
| No | 56 (57.1) | 24 (60.0) | 80 (58.0) | — |
| **Diabetes mellitus, n (%)** | | | | |
| Yes | 13 (13.3) | 7 (17.5) | 20 (14.5) | .52[a] |
| No | 85 (86.7) | 33 (82.5) | 118 (85.5) | — |
| **Hypertension, n (%)** | | | | |
| Yes | 29 (29.6) | 13 (32.5) | 42 (30.4) | .74[a] |
| No | 69 (70.4) | 27 (67.5) | 96 (69.6) | — |

[a]Comparison of the differences are given according to the Pearson chi-square test (statistical significance at $P<.05$).

[b]Not applicable.

[c]Comparison of the differences are given according to the Fisher exact test (statistical significance at $P<.05$).

**Table 5.** Lipid profile, ApoBa concentrations, and other indices stratified by allele.

| Parameters | Allele | | | *P* value |
|---|---|---|---|---|
| | A (n=98) | G (n=40) | Total (N=138) | |
| **TC[b] (mmol/L)** | | | | |
| Normal, n (%) | 3 (3.1) | 5 (12.5) | 8 (5.8) | .10[c] |
| Borderline high, n (%) | 17 (17.3) | 7 (17.5) | 24 (17.4) | —[d] |
| High, n (%) | 78 (79.6) | 28 (70) | 106 (76.8) | — |
| Mean (SD) | 7.05 (1.2) | 6.74 (1.28) | 6.96 (1.22) | .18[e] |
| **TGs[f] (mmol/L)** | | | | |
| Normal, n (%) | 31 (31.6) | 15 (37.5) | 46 (33.3) | .07[c] |
| Borderline high, n (%) | 67 (68.4) | 23 (57.5) | 90 (65.2) | — |
| High, n (%) | 0 (0) | 2 (5) | 2 (1.4) | — |
| Mean (SD) | 2.45 (0.65) | 2.51 (0.93) | 2.47 (0.74) | .68[e] |
| **HDL-C[g] (mmol/L)** | | | | |
| Normal, n (%) | 84 (85.7) | 38 (95) | 122 (88.4) | .15[c] |
| Decreased, n (%) | 14 (14.3) | 2 (5) | 16 (11.6) | — |
| Mean (SD) | 1.31 (0.35) | 1.46 (0.39) | 1.36 (0.37) | .05[e] |
| **LDL-C[h] (mmol/L)** | | | | |
| Borderline high, n (%) | 22 (22.4) | 18 (45.0) | 40 (29.0) | .008[c] |
| High, n (%) | 76 (77.6) | 22 (55.0) | 98 (71.0) | — |
| Mean (SD) | 4.91 (1.02) | 4.36 (0.97) | 4.75 (1.03) | .004[e] |
| Non–HDL-C (mmol/L), mean (SD) | 5.73 (1.14) | 5.28 (1.24) | 5.6 (1.18) | .04[e] |
| ApoB (mg/dL), mean (SD) | 145.56 (23.23) | 135.88 (15.99) | 142.75 (21.78) | .02[e] |
| Hemoglobin (g/dL), mean (SD) | 13.82 (1.28) | 13.7 (1.5) | 13.78 (1.34) | .63[e] |
| HbA1c[i] (%), mean (SD) | 6.1 (1.75) | 6.29 (1.82) | 6.16 (1.76) | .58[e] |
| Glucose (mmol/L), mean (SD) | 5.87 (1.83) | 5.96 (1.84) | 5.9 (1.83) | .80[e] |
| Creatinine (µmol/L), mean (SD) | 71.81 (15.83) | 69.85 (15.1) | 71.24 (15.59) | .51[e] |
| Urea (mmol/L), mean (SD) | 5.2 (1.66) | 5.03 (1.55) | 5.15 (1.62) | .58[e] |

[a]ApoB: apolipoprotein B.

[b]TC: total cholesterol.

[c]Comparison of the differences are given according to the Fisher-Freeman-Halton exact test (statistical significance at $P<.05$).

[d]Not applicable.

[e]Comparison of the differences are given according to the independent samples test (statistical significance at $P<.05$).

[f]TG: triglyceride.

[g]HDL-C: high-density lipoprotein cholesterol.

[h]LDL-C: low-density lipoprotein cholesterol.

[i]HbA1c: hemoglobin A1c.

The violin plots also demonstrated that carriers of allele A exhibit higher concentrations of both LDL-C and ApoB compared to carriers of the G allele (Figure 4).

After adjusting for age, sex, the BMI, and diabetes, the *APOB* rs676210 variant remained significantly associated with increased LDL-C, non–HDL-C, and ApoB levels, as shown in Table 6. The association with HDL-C, however, was attenuated and no longer statistically significant. These results suggest a robust link between the rs676210 genotype and atherogenic lipid parameters, independent of major metabolic confounders.

**Figure 4.** Violin plots of LDL-C and ApoB concentrations by allele. ApoB: apolipoprotein B; FDR-P: false discovery rate–adjusted *P* value; LDL-C: low-density lipoprotein cholesterol.

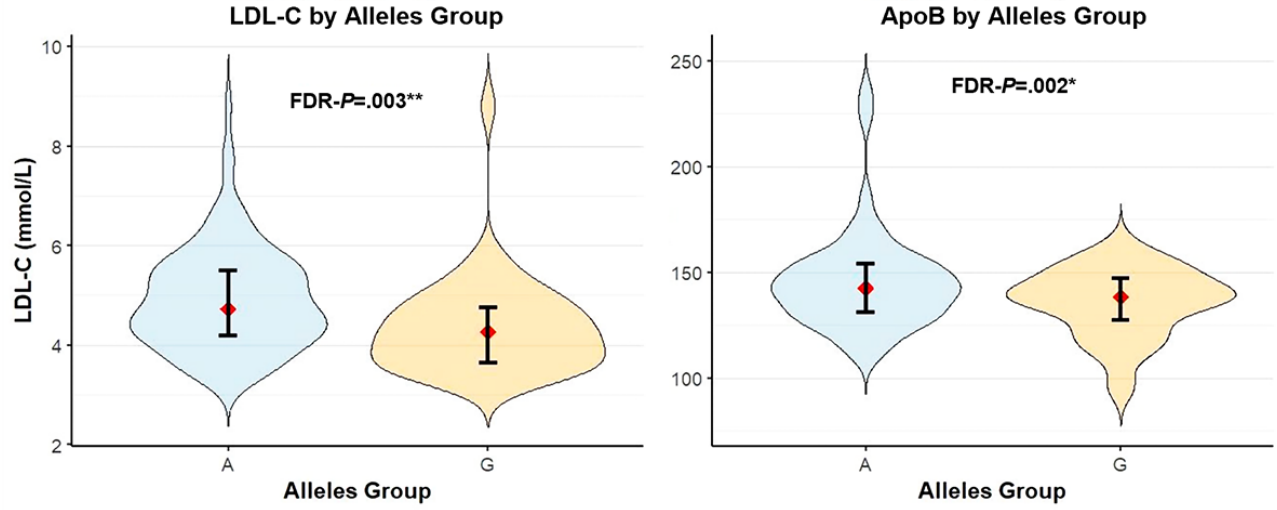



**Table 6.** Association between *APOB*[a] rs676210 alleles and lipid profile outcomes: crude and adjusted estimates with multivariable linear regression.

| Outcome | Crude estimate *P* value[b] | Model 1[c] adjusted β (SE; 95% CI) | Model 1 *P* value | Model 2[d] adjusted β (SE; 95% CI) | Model 2 *P* value | FDR-*P*[e] |
|---|---|---|---|---|---|---|
| HDL-C[f] | .05 | –0.091 (0.067; (–0.224 to 0.041) | .17 | –0.075 (0.067; (–0.207 to 0.057) | .26 | .18 |
| LDL-C[g] | .004 | 0.606 (0.196; (0.218 to 0.993) | .002 | 0.621 (0.198; (0.230 to 1.012) | .002 | .01 |
| Non–HDL-C | .04 | 0.512 (0.228; (0.061 to 0.963) | .03 | 0.549 (0.229; (0.097 to 1.001) | .02 | .05 |
| ApoB | .02 | 7.601 (4.092; (–0.494 to 15.695) | .07 | 8.202 (4.109; (0.075 to 16.330) | .05 | .09 |

[a]*APOB*: apolipoprotein B.

[b]Crude *P* values from unadjusted comparisons between genotype groups.

[c]Model 1: age, sex, and BMI.

[d]Model 2: age, sex, BMI, and diabetes.

[e]FDR-P: false discovery rate–adjusted *P* value calculated for model 1 comparisons to account for multiple testing.

[f]HDL-C: high-density lipoprotein cholesterol.

[g]LDL-C: low-density lipoprotein cholesterol.

# Discussion

## Principal Findings

In this Vietnamese cohort with newly diagnosed hyper–LDL-C, we observed that individuals carrying the rs676210 AA genotype had a markedly more atherogenic lipid profile than G allele carriers (GA or GG). Specifically, the AA genotype was associated with significantly higher LDL-C, non–HDL-C, and ApoB levels, alongside lower HDL-C. In addition, allele-based analysis revealed that carriers of allele A also had higher LDL-C, non–HDL-C, ApoB levels, and lower HDL-C than G allele carriers.

This pattern suggests that the A allele of rs676210 may predispose to the accumulation of ApoB-containing lipoproteins in circulation. The biological basis for this association likely stems from the functional role of the *APOB* gene variant. Rs676210 causes a Pro2739Leu substitution in the ApoB-100 protein, a change that can alter the protein's conformation and interactions [6]. Proline is a rigid, helix-breaking residue, whereas leucine is a hydrophobic, helix-forming amino acid [26]; replacing proline with leucine at position 2739 could conceivably influence how ApoB-100 folds or binds to lipid and receptor molecules [27,28]. One consequence of this substitution, as reported in prior studies, is an effect on the susceptibility of LDL particles to oxidation [11,29]. A genome-wide study pinpointed rs676210 as a regulator of oxidized LDL levels, which is noteworthy because oxidized LDL is highly atherogenic and plays a key role in triggering foam cell formation and early atherosclerotic lesions [8]. It is possible that the leucine-encoding A allele produces LDL particles that are less prone to oxidative modification, as one earlier report suggests this variant renders LDL "less" easily oxidized [8]. If LDL is less readily oxidized, it might evade rapid uptake by macrophages and persist longer in circulation, contributing to higher measured LDL-C and ApoB levels. Conversely, the G allele (proline variant) could make LDL particles more susceptible to oxidation and clearance, potentially resulting in relatively lower LDL-C but more oxidative stress per particle. This hypothesis aligns with the notion that rs676210

is “functional” in modifying LDL particle behavior. Of note, our finding that A allele homozygotes have lower HDL-C also points to a broader dysregulation of lipid metabolism associated with this variant, though the mechanism for the HDL effect is unclear. It may be secondary to the remodeling of lipoproteins in an environment of high ApoB lipoprotein concentration. Further biochemical studies are needed to delineate how the Pro2739Leu substitution influences LDL receptor binding, particle clearance rates, or hepatic lipid homeostasis. Nonetheless, the established link between this SNP and LDL oxidation provides a plausible mechanistic bridge from the genotype to the pro-atherogenic lipid phenotype observed in our subjects [7,8,30,31].

The relationship between rs676210 and lipid levels has been examined in several populations, and our results both corroborate and diverge from prior findings. Interestingly, the direction of association observed in our Vietnamese cohort (A allele associated with higher LDL-C and ApoB and lower HDL-C) contrasts with some reports in European ancestry studies. Chasman et al [9], in an extensive genome-wide analysis, noted that the A allele of rs676210 is linked to lower LDL-C and TGs and higher HDL-C. This initially counterintuitive discrepancy highlights the complexity of gene effect modulation by the ethnic and environmental context. In a Chinese population study [11], which focused on myocardial infarction (MI) risk, the G allele of rs676210 (coding for proline at 2739) was identified as the risk variant: Chinese individuals carrying the G allele had higher plasma ApoB levels and an increased risk of MI [11]. This Chinese study also observed a trend toward higher LDL-C in G carriers, although it did not reach statistical significance. These findings imply that in East Asian populations, the G allele may be deleterious, whereas the A (leucine) allele might be relatively protective, consistent with the direction reported by Chasman et al [9] in predominantly European cohorts. By contrast, our findings align more closely with those from a Western Mexican population. Aceves-Ramírez et al [7] reported that Mexican individuals with the AA genotype of rs676210 have significantly elevated odds of acute coronary syndrome, and overall, the A allele confers a higher risk of coronary events compared to allele G (odds ratio [OR] 1.72, $P<.001$). In addition, the A allele frequency was lower in controls (22.5%) than in cases (33%), suggesting the A variant is the risk allele in this population [7]. This parallels our Vietnamese cohort results, where A allele carriers showed a worse lipid profile, consistent with a risk-promoting effect. Notably, allele frequency and linkage disequilibrium patterns for rs676210 vary among ethnic groups. Specifically, East Asians (eg, Han Chinese) have been reported to have a higher A allele frequency (the putatively “normal” allele in those groups) [11], whereas in populations of European or mixed ancestry, the A allele may be the minor variant. Such differences could lead to flip-flopping of which allele appears as “risk” in genetic association studies due to interactions with other genetic loci (epistasis) or environmental factors. Additionally, the context of the study cohorts differs. The Chinese study involved unselected patients with MI and controls [11], whereas ours focused on individuals already flagged for high LDL-C. The latter might represent a subset enriched for genetic hyperlipidemia traits, potentially amplifying the impact of specific alleles. In a population under strong dietary influences or with different baseline risk factor profiles, the effect of rs676210 on measured lipids could manifest differently. For example, high-carbohydrate diets standard in parts of Asia [32-34] might modulate TG-rich VLDL production, interacting with *APOB* variants to influence LDL composition [35,36]. Although speculative, such gene-environment interplay could partly explain why the same SNP shows heterogeneous associations across studies.

Despite these discrepancies, all studies, including ours, reinforce that rs676210 is not a neutral polymorphism but one that influences lipid metabolism in some fashion. Whether the A allele is beneficial or detrimental may depend on the metabolic context. Some have proposed that the leucine variant (A allele) might produce LDL particles less prone to oxidation (potentially reducing inflammatory risk). However, if those particles circulate longer, they could raise LDL-C levels—a trade-off between the quantity and quality of LDL. In contrast, the proline variant (G allele) might shorten LDL residence time at the cost of being more atherogenic per particle. More research is required to resolve these complex dynamics, as well as exceptionally functional assays and population-specific analyses. Our findings contribute to this dialogue by providing data from a Southeast Asian cohort, illustrating that the rs676210-lipid association may parallel that seen in specific Western populations (risk allele A) rather than the pattern reported in East Asians (risk allele G). This underscores the importance of investigating genetic associations within diverse ethnic groups rather than extrapolating findings universally. It should also be noted that participants in our study were identified through annual occupational health examinations, and none had received prior lipid-lowering therapy. Although this allows for the characterization of genotype-phenotype associations in treatment-naive individuals, it may limit generalizability to populations with established CVD or ongoing lipid management.

## Limitations

Our study should be clarified with strengths and limitations. From a clinical and public health perspective, identifying a significant association between *APOB* rs676210 and lipid profiles in Vietnamese patients may carry exploratory implications. First, it highlights a potential genetic marker that could be used to refine risk stratification for dyslipidemia and its downstream consequences. In settings where comprehensive genomic screening is not feasible, testing for a limited panel of impactful SNPs, such as rs676210, could serve as an exploratory tool to flag and identify individuals with a heritable propensity for elevated ApoB and LDL-C. For example, if an individual is known to carry the AA genotype of rs676210, our cross-sectional findings indicate a trend toward a more adverse lipid profile (high LDL-C, high ApoB). These findings may help generate hypotheses regarding whether carriers of the AA genotype are at increased risk for premature ASCVD and could potentially benefit from earlier or more intensive intervention strategies. However, clinical applications remain speculative and would require validation in prospective or interventional studies. Moreover, our sample was drawn from a hospital-based cohort identified through routine occupational health screenings, which may not be fully representative of the general population.

Specifically, because all participants were recruited from a hospital-based setting and selected based on newly diagnosed elevated LDL-C without prior statin use, our sample is inherently enriched for individuals with more clinically overt dyslipidemia. This may have amplified the observed genotype-phenotype associations. When compared with findings from 2 recent Vietnamese population-based surveys, the lipid profiles in our cohort appear markedly more atherogenic, underscoring the selection bias and limiting generalizability to the broader community. Caution is therefore warranted when extrapolating these findings beyond the clinical screening context. This is particularly relevant in Vietnam and similar developing contexts, where ASCVD often presents at a relatively young age, and resource-intensive interventions should be targeted to those at most significant risk. Another important implication of our study is the potential for personalized therapy. Evidence suggests that genetic polymorphisms in lipid metabolism genes can influence treatment response. Interestingly, the rs676210 variant has been studied in pharmacogenetic contexts: an Iraqi study demonstrated that patients with the AA genotype experience a greater LDL-C reduction in response to high-dose atorvastatin therapy [6]. Although our study did not involve a treatment intervention, the observed association between the AA genotype and elevated LDL-C levels raises the possibility that individuals carrying this high-risk genotype might benefit more from intensive lipid-lowering therapy, such as high-dose statins. However, this interpretation remains speculative, as our study was not designed to evaluate pharmacogenetic responses. To determine whether the rs676210 polymorphism predicts differential response to statins, a prospective, randomized controlled trial or a genotype-stratified cohort study with pre- and posttreatment lipid measurements would be required. Such a study would help assess whether the AA genotype is not only a marker of increased risk but also a predictor of treatment efficacy. From a hypothesis-generating standpoint, carriers of the A allele might represent a subgroup warranting closer monitoring and, potentially, more intensive therapy—pending validation in longitudinal or interventional studies. It is also worth noting the novelty of our findings. To the best of our knowledge, this is the first report detailing the association of rs676210 with lipid profiles in a Vietnamese cohort. The Vietnamese population has a distinct genetic architecture due to its ethnic background and has been underrepresented in genomic research. Our study contributes preliminary data suggesting that a common *APOB* variant studied in other populations may also be relevant in the Vietnamese context, though confirmation is needed in larger cohorts. Additionally, it is important to note that comprehensive genotype frequency data for the Vietnamese population remain scarce, particularly for variants related to lipid metabolism, such as *APOB* rs676210. This lack of large-scale population-based genetic studies limited our ability to compare the genotype distribution observed in our sample with national reference values or those of neighboring countries. As such, our findings should be interpreted as exploratory and hypothesis generating, reinforcing the need for broader genomic research in this underrepresented population. Lastly, we acknowledge that no sensitivity or subgroup analyses were prespecified, and therefore the findings should be interpreted with appropriate caution.

## Conclusion

In conclusion, this study suggests that the *APOB* rs676210 polymorphism is associated with lipid profile parameters, including LDL-C, non–HDL-C, HDL-C, and ApoB levels, in a Vietnamese population newly diagnosed with elevated LDL-C. Although the study has limitations, including potential selection bias, a moderate sample size, and lack of functional measures (eg, oxidized LDL levels), our findings are biologically coherent and supported by external studies, enhancing confidence in the validity and potential clinical relevance of the association.

## Acknowledgments

We would like to express our sincere gratitude to the Can Tho University of Medicine and Pharmacy and the Can Tho University of Medicine and Pharmacy Hospital for their invaluable support and provision of essential resources that made this study possible.

## Data Availability

The datasets generated and analyzed during this study are available in the Gen repository on GitHub [37].

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Hardy-Weinberg equilibrium (HWE) for the apolipoprotein B (*APOB*) rs676210 variant.
[DOCX File , 29 KB - cardio_v9i1e76850_app1.docx ]

Multimedia Appendix 2
Data collection form (English version).
[PDF File (Adobe PDF File), 147 KB - cardio_v9i1e76850_app2.pdf ]

Multimedia Appendix 3

Data collection form (Vietnamese version).
[PDF File (Adobe PDF File), 173 KB - cardio_v9i1e76850_app3.pdf ]

## References


1. Ferhatbegović L, Mršić D, Kušljugić S, Pojskić B. LDL-C: only causal risk factor for ASCVD. why is it still overlooked and underestimated? Curr Atheroscler Rep 2022 Aug 28;24(8):635-642. [doi: 10.1007/s11883-022-01037-3] [Medline: 35635632]
2. Lee Z, Llanes EJ, Sukmawan R, Thongtang N, Ho HQT, Barter P, Cardiovascular RISk Prevention (CRISP) in Asia Network. Prevalence of plasma lipid disorders with an emphasis on LDL cholesterol in selected countries in the Asia-Pacific region. Lipids Health Dis 2021 Apr 15;20(1):33 [FREE Full text] [doi: 10.1186/s12944-021-01450-8] [Medline: 33858442]
3. Ngoc Pham NT, Thao Tran LN, Trung CH, Phan HM, Bui DT, Tran CM, et al. Prognostic value of polymorphism in the gene for coronary artery lesions and cardiovascular events in acute myocardial infarction patients. Int J Cardiol Cardiovasc Risk Prev 2025 Sep;26:200458 [FREE Full text] [doi: 10.1016/j.ijcrp.2025.200458] [Medline: 40636316]
4. Tuan Huynh A, Vu HA, Chuong HQ, Anh TH, Viet Tran A. The rs243865 polymorphism in matrix metalloproteinase-2 and its association with target organ damage in patients with resistant hypertension: cross-sectional study. JMIR Cardio 2025 May 01;9:e71016-e71016 [FREE Full text] [doi: 10.2196/71016] [Medline: 40311090]
5. Tran DC, Le LHG, Thai TT, Van Hoang S, Do MD, Truong BQ. Effect of AGTR1 A1166C genetic polymorphism on coronary artery lesions and mortality in patients with acute myocardial infarction. PLoS One 2024 Apr 18;19(4):e0300273 [FREE Full text] [doi: 10.1371/journal.pone.0300273] [Medline: 38635772]
6. Abdulfattah SY, Al-Awadi SJ. ApoB gene polymorphism (rs676210) and its pharmacogenetics impact on atorvastatin response among Iraqi population with coronary artery disease. J Genet Eng Biotechnol 2021 Jun 22;19(1):95 [FREE Full text] [doi: 10.1186/s43141-021-00193-4] [Medline: 34156559]
7. Aceves-Ramírez M, Valle Y, Casillas-Muñoz F, Martínez-Fernández DE, Parra-Reyna B, López-Moreno VA, et al. Analysis of the APOB gene and apolipoprotein B serum levels in a Mexican population with acute coronary syndrome: association with the single nucleotide variants rs1469513, rs673548, rs676210, and rs1042034. Genet Res (Camb) 2022;2022:4901090 [FREE Full text] [doi: 10.1155/2022/4901090] [Medline: 35440891]
8. Mäkelä K, Seppälä I, Hernesniemi JA, Lyytikäinen L, Oksala N, Kleber ME, et al. Genome-wide association study pinpoints a new functional apolipoprotein B variant influencing oxidized low-density lipoprotein levels but not cardiovascular events. Circ Cardiovasc Genet 2013 Feb;6(1):73-81. [doi: 10.1161/circgenetics.112.964965]
9. Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. PLoS Genet 2009 Nov 20;5(11):e1000730 [FREE Full text] [doi: 10.1371/journal.pgen.1000730] [Medline: 19936222]
10. Gu Q, Han Y, Lan Y, Li Y, Kou W, Zhou Y, et al. Association between polymorphisms in the APOB gene and hyperlipidemia in the Chinese Yugur population. Braz J Med Biol Res 2017;50(11):e6613. [doi: 10.1590/1414-431x20176613]
11. Liu C, Yang J, Han W, Zhang Q, Shang X, Li X, et al. Polymorphisms in ApoB gene are associated with risk of myocardial infarction and serum ApoB levels in a Chinese population. Int J Clin Exp Med 2015;8(9):16571-16577 [FREE Full text] [Medline: 26629186]
12. Helmy M, Awad M, Mosa KA. Limited resources of genome sequencing in developing countries: challenges and solutions. Appl Transl Genom 2016 Jun;9:15-19 [FREE Full text] [doi: 10.1016/j.atg.2016.03.003] [Medline: 27354935]
13. Dang AK, Thi Le LT, Pham NM, Nguyen DQ, Thi Nguyen HT, Dang SC, et al. An upward trend of dyslipidemia among adult population in Vietnam: evidence from a systematic review and meta-analysis. Diabetes Metab Syndr 2025 Jan;19(1):103171 [FREE Full text] [doi: 10.1016/j.dsx.2024.103171] [Medline: 39700840]
14. Vo NX, Pham HL, Bui TT, Bui TT. Systematic review on efficacy, effectiveness, and safety of pitavastatin in dyslipidemia in Asia. Healthcare (Basel) 2024 Dec 31;13(1):59 [FREE Full text] [doi: 10.3390/healthcare13010059] [Medline: 39791666]
15. Tran AV, Nguyen TT, Tran LNT, Nguyen PM, Nguyen T. Efficacy of rosuvastatin and atorvastatin in Vietnamese patients with acute coronary syndrome: a randomized trial. Pharm Sci Asia 2021;48(5):413-419. [doi: 10.29090/psa.2021.05.20.144]
16. Tsao C, Aday A, Almarzooq Z, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation 2022 Feb 22;145(8):e153-e639. [doi: 10.1161/cir.0000000000001052]
17. Civeira F, Arca M, Cenarro A, Hegele RA. A mechanism-based operational definition and classification of hypercholesterolemia. J Clin Lipidol 2022 Nov;16(6):813-821 [FREE Full text] [doi: 10.1016/j.jacl.2022.09.006] [Medline: 36229375]
18. Zhang Y, An J, Reynolds K, Safford MM, Muntner P, Moran AE. Trends of elevated low-density lipoprotein cholesterol, awareness, and screening among young adults in the US, 2003-2020. JAMA Cardiol 2022 Oct 01;7(10):1079-1080 [FREE Full text] [doi: 10.1001/jamacardio.2022.2641] [Medline: 36001331]
19. Weir CB, Jan A. BMI Classification Percentile And Cut Off Points. Treasure Island, FL: StatPearls Publishing LLC; 2025.
20. Mancia G, Kreutz R, Brunström M. 2023 ESH guidelines for the management of arterial hypertension. The Task Force for the management of arterial hypertension of the European Society of Hypertension: endorsed by the International Society

of Hypertension (ISH) and the European Renal Association (ERA). J Hypertens 2023;41(12):1874-2071. [doi: 10.1097/hjh.0000000000003621]
21. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 International Society of Hypertension global hypertension practice guidelines. Hypertension 2020 Jun;75(6):1334-1357. [doi: 10.1161/hypertensionaha.120.15026]
22. Salazar JH. Overview of urea and creatinine. Lab Med 2014 Feb 01;45(1):e19-e20. [doi: 10.1309/lm920sbnzpjrjgut]
23. ElSayed N, Aleppo G, Aroda V, Bannuru RR, Brown FM, Bruemmer D, on behalf of the American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of care in diabetes-2023. Diabetes Care 2023 Jan 01;46(Suppl 1):S19-S40 [FREE Full text] [doi: 10.2337/dc23-S002] [Medline: 36507649]
24. Friedewald W, Levy R, Fredrickson D. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18(6):499-502. [doi: 10.1093/clinchem/18.6.499]
25. R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2024.
26. Breimann S, Kamp F, Steiner H, Frishman D. AAontology: an ontology of amino acid scales for interpretable machine learning. J Mol Biol 2024 Oct 01;436(19):168717 [FREE Full text] [doi: 10.1016/j.jmb.2024.168717] [Medline: 39053689]
27. Wojczynski MK, Gao G, Borecki I, Hopkins PN, Parnell L, Lai C, et al. Apolipoprotein B genetic variants modify the response to fenofibrate: a GOLDN study. J Lipid Res 2010 Nov;51(11):3316-3323. [doi: 10.1194/jlr.p001834]
28. Weber C, Noels H. Atherosclerosis: current pathogenesis and therapeutic options. Nat Med 2011 Nov 07;17(11):1410-1422. [doi: 10.1038/nm.2538] [Medline: 22064431]
29. Mäkelä KM, Traylor M, Oksala N, Kleber ME, Seppälä I, Lyytikäinen LP, Wellcome Trust Case Control Consortium 2 (WTCCC2), et al. Association of the novel single-nucleotide polymorphism which increases oxidized low-density lipoprotein levels with cerebrovascular disease events. Atherosclerosis 2014 May;234(1):214-217. [doi: 10.1016/j.atherosclerosis.2014.03.002] [Medline: 24681816]
30. Benn M, Stene M, Nordestgaard B, Jensen GB, Steffensen R, Tybjaerg-Hansen A. Common and rare alleles in apolipoprotein B contribute to plasma levels of low-density lipoprotein cholesterol in the general population. J Clin Endocrinol Metab 2008 Mar;93(3):1038-1045. [doi: 10.1210/jc.2007-1365] [Medline: 18160469]
31. Teslovich T, Musunuru K, Smith A, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010 Aug 05;466(7307):707-713 [FREE Full text] [doi: 10.1038/nature09270] [Medline: 20686565]
32. Harris J, Nguyen PH, Tran LM, Huynh PN. Nutrition transition in Vietnam: changing food supply, food prices, household expenditure, diet and nutrition outcomes. Food Sec 2020 Sep 01;12(5):1141-1155. [doi: 10.1007/s12571-020-01096-x]
33. Lim CG, Tai ES, van Dam RM. Replacing dietary carbohydrates and refined grains with different alternatives and risk of cardiovascular diseases in a multi-ethnic Asian population. Am J Clin Nutr 2022 Mar 04;115(3):854-863 [FREE Full text] [doi: 10.1093/ajcn/nqab403] [Medline: 34996115]
34. Choi H, Song S, Kim J, Chung J, Yoon J, Paik H, et al. High carbohydrate intake was inversely associated with high-density lipoprotein cholesterol among Korean adults. Nutr Res 2012 Feb;32(2):100-106. [doi: 10.1016/j.nutres.2011.12.013] [Medline: 22348458]
35. Witztum JL, Schonfeld G. Carbohydrate diet-induced changes in very low density lipoprotein composition and structure. Diabetes 1978 Dec 01;27(12):1215-1229. [doi: 10.2337/diabetes.27.12.1215]
36. Borén J, Taskinen M, Packard C. Biosynthesis and metabolism of ApoB-containing lipoproteins. Annu Rev Nutr 2024 Aug;44(1):179-204 [FREE Full text] [doi: 10.1146/annurev-nutr-062222-020716] [Medline: 38635875]
37. DragonRanSet/Gen. GitHub. URL: https://github.com/DragonRanSet/Gen [accessed 2025-07-16]

## Abbreviations

**ApoB/*APOB*:** apolipoprotein B
**ASCVD:** atherosclerotic cardiovascular disease
**CVD:** cardiovascular diseaseHbA1c: hemoglobin A1c
**HDL-C:** high-density lipoprotein cholesterol
**HEX:** hexachloro-fluorescein
**FAM:** 6-carboxyfluorescein
**LDL:** low-density lipoprotein
**LDL-C:** low-density lipoprotein cholesterol
**MI:** myocardial infarction
**PCR:** polymerase chain reaction
**SNP:** single nucleotide polymorphism
**TC:** total cholesterol
**TG:** triglyceride
**VLDL:** very low-density lipoprotein
**WB:** Wash Buffer

*Edited by A Coristine; submitted 02.05.25; peer-reviewed by X Ding, RS Gomaa Mahmoud, H Tien; comments to author 20.05.25; revised version received 16.06.25; accepted 09.07.25; published 07.08.25.*

*Please cite as:*
*Nguyen QT, Tran AV, Nguyen BT, Nguyen HT, Thai NTH, Phan HH*
*Lipid Profile and Apolipoprotein B Serum Levels in the Vietnamese Population With Newly Diagnosed Elevated Low-Density Lipoprotein Cholesterol and Association With the Single-Nucleotide Variant rs676210: Cross-Sectional Study*
*JMIR Cardio 2025;9:e76850*
*URL: https://cardio.jmir.org/2025/1/e76850*
*doi:10.2196/76850*
*PMID:40773287*

# Using Patient-Held Devices to Measure Variations in Resting Heart Rate and Step Count Prior to Presentation With an Acute Illness: International, Multicenter Flash Mob Feasibility Study

Jason G A den Duijn[1], MSc; Ahmed A M Hajjaj[2], MD; John Kellett[3], MD; Erika Frischknecht Christensen[4], MD, PhD; Harm R Haak[5,6], MD, PhD; Mikkel Brabrand[7], MD, PhD; Christian H Nickel[8], MD, PhD; Prabath W B Nanayakkara[9], MD, PhD; Christian P Subbe[2,10], MD; Jelmer Alsma[11], MD, PhD; Safer@Home[12,13]; Research Consoritum Acute Medicine (ORCA)[12,14]

[1]Department of Medical Oncology, Erasmus University Medical Center, Dr. Molewaterplein 40, Rotterdam, The Netherlands

[10]North Wales Medical School, Bangor University, Bangor, The Netherlands

[11]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands

[12]See Acknowledgments

13

14

[2]Department of Acute Medicine, Ysbyty Gwynedd Hospital National Health Service Trust, Bangor, United Kingdom

[3]School of Clinical and Biomedical Sciences, University of Greater Manchester, Bolton, United Kingdom

[4]Center for Prehospital and Emergency Research, Department of Emergency Medicine and Trauma Center, Aalborg Hospital and Aalborg University, Aalborg, Denmark

[5]Department of Internal Medicine, Máxima Medisch Centrum, Veldhoven, The Netherlands

[6]Department of Health Services Research, Aging and Long Term Care, Care and Public Health Research Institute School for Public Health and Primary Care, Maastricht, The Netherlands

[7]Department of Emergency Medicine, Odense University Hospital, Odense, Denmark

[8]Emergency Department, University Hospital Basel, University of Basel, Basel, Switzerland

[9]Section of Acute Medicine, Department of Internal Medicine, Amsterdam Public Health Research Institute, Amsterdam University Medical Center Location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

**Corresponding Author:**
Jason G A den Duijn, MSc
Department of Medical Oncology, Erasmus University Medical Center, Dr. Molewaterplein 40, Rotterdam, The Netherlands

## Abstract

**Background:** Many patients experience a gradual decline in health before seeking hospital care, with subtle changes in vital signs such as increased heart rate or decreased mobility. Recognizing deviations from baseline vital signs can support clinical decision-making, especially admission decisions. Smart devices (ie, smartphones, smartwatches, and activity trackers) track health metrics like heart rate and step count, offering new opportunities to estimate illness severity and track deterioration early.

**Objective:** This study aimed to assess the feasibility of using heart rate and step count measurements from smart devices (ie, smartphones, smartwatches, and activity trackers) to enhance the evaluation of patients presenting with acute illness in emergency settings.

**Methods:** We conducted an international multicenter prospective observational study using the flash mob study design in 34 hospitals in the Netherlands (n=17), the United Kingdom (n=7), Denmark (n=9), and Switzerland (n=1) in May 2024. Researchers collaborated with patients to complete questionnaires upon an acute care (ie, emergency department, acute medical unit, same day emergency care) visit and extracted physiological data from their smart devices.

**Results:** Among patients with an acute care visit (n=1137), 40% (n=452) had a smart device with health data. These patients tended to be from a higher educational level and in relatively good health. Only half had retrievable heart rate or step count data, resulting in a usable data set for 20% (n=209) of the total study population. Analysis showed a significant increase in heart rate ($P$<.001) and a decrease in step count ($P$<.001) in the days preceding their hospital visit. Both heart rate ($P$=.04) and step count ($P$=.04) on the day before presentation were significantly associated with disposition.

**Conclusions:** Our study demonstrates the feasibility of using a patient's personal smart device to monitor vital signs in the days leading up to an acute care visit. In a selected patient group, significant changes in heart rate and step count were observed prior

to hospital presentation, suggesting that disposition may be predicted using data collected from the patient's own device. High-risk patient groups, who might benefit the most from digital health monitoring, are currently underrepresented among device users.

***(JMIR Cardio 2025;9:e76218)*** doi:10.2196/76218

**KEYWORDS**

smart device; acute care; flash mob; vital signs; monitoring

## Introduction

Many patients presenting to hospitals with acute medical complaints experience a gradual decline in health over days or even weeks before seeking care [1]. During this period, subtle yet significant changes in vital signs, such as an increased heart rate or decreased step count, may already be present, signaling underlying physiological stress [2-4]. In the emergency department (ED), acute medical unit (AMU), or same-day emergency care (SDEC), the decision to admit a patient to the hospital is influenced by various factors, including patient symptoms, vital signs, past medical history, medication use, diagnosis, as well as availability of social support at home [5-7], and the estimated risk of clinical deterioration in the subsequent hours and days [8]. Assessing how far a patient's vital signs deviate from their baseline may provide valuable insight to support admission decisions.

Reference values from healthy individuals, recorded during periods of physiological stability, are often used for this purpose [9]. The extent of deviation from physiological baseline, as well as the number of vital signs affected, correlates positively with the frequency of adverse events and is widely recognized as a measure of illness severity [9]. These above or below average "abnormalities" can be assessed with generic tools such as the National Early Warning Score (NEWS) [10] or with disease-specific tools such as the CURB65 for pneumonia [11] and the Blatchford score for bleeding in the upper gastrointestinal tract [12]. These scoring systems are based on population data and may under- or overestimate risk in individual patients. The vital signs of a healthy individual have a natural variability influenced by several factors, including age, sex, body composition, medication use, genetics, and physical condition [13-15]; very fit individuals, for example, often have very low heart rates [16]. Therefore, personal baseline measurements may offer a more tailored approach to evaluating deviations than population averages to determine the illness severity of an individual patient [17].

The use of smart devices is growing rapidly, with reports of 76% to 97% of the population in the United States [18] and 82% to 98% in the United Kingdom [19] owning smartphones, with rates varying between age groups. Many smart devices (ie, smartphones, smartwatches, and activity trackers) measure basic health metrics, including heart rate and step count. These may provide an opportunity for patients to track and share their baseline values over time [20], potentially enabling a more accurate estimation of their illness severity [1] and its rate of deterioration [21]. This has already been demonstrated for several medical conditions, for example, new or paroxysmal atrial fibrillation and COVID-19 [22-27].

The aim of this study was to assess the feasibility of using heart rate and step count measurements captured by smart devices (ie, smartphones, smartwatches, and activity trackers) to enhance the assessment of patients presenting as emergencies with an acute illness. Feasibility was defined as having a smart device with heart rate or step count data recorded at least twice between 30 days (baseline) before and on the day of admission, enabling trend analysis, and measured as the percentage of patients with an acute care visit (ED, AMU, or SDEC).

## Methods

### Study Design and Setting

We conducted an international, multicenter, prospective observational study using the flash mob study design [28] across 34 hospitals in four countries in May 2024: the Netherlands (n=17), the United Kingdom (n=7), Denmark (n=9), and Switzerland (n=1). This study was previously described in a protocol paper [29]. Each site recruited patients over a single day between 8 AM and 10 PM.

### Ethical Considerations

Ethical approval was obtained from the Medical Ethical Assessment Committee (MECC-2022-0795) of the Erasmus University Medical Center, Rotterdam, the Netherlands, and the London–Harrow Research Ethics Committee, Bristol, United Kingdom (IRAS 321129). Approval for all other study sites was acquired in accordance with national and local regulations. All patients gave written informed consent to participate in the study. All collected data were pseudonymized before data analysis.

### Participants

All patients presenting to ED, AMU, or SDEC were screened for inclusion and exclusion criteria and asked about their use of monitoring devices. Inclusion criteria were ≥18 years, having a device (smartphone or smartwatch) capable of measuring resting heart rate or step count, and the ability to provide informed consent. Patients were excluded for presentation with trauma or clinical instability, as determined by the treating physician. Written informed consent was obtained by trained investigators. Researchers collaborated with patients to complete questionnaires, after which physiological data were extracted from the patients' devices.

### Variables

Patient variables included gender, age groups (18 - 30, 31 - 50, 51 - 65, and 65+ years), educational attainment, digital literacy, and device brand and type. The resting, maximum, and minimum heart rates were collected from the patient's device for 4 time points: 30 days, 7 days, 1 day before the hospital

visit, and on the day of admission. In addition, the heart rate measured at presentation to hospital was extracted from the patient's clinical record. Step count data were collected from the patient's device for 9 time periods: 30 days, the preceding 7 days before the hospital visit, and the day of the visit. As changes in step count fluctuate more, we chose to collect these on more time points than heart rate. For either heart rate or step count, a minimum of two time points had to be available for data analysis. Additionally, the value of the NEWS, the Clinical Frailty Scale (CFS) [30], and a standardized assessment of gait were recorded. Follow-up data collected up to 7 days after presentation included the date of visit, disposition (admission to hospital or discharge), admission to intensive care or high care areas, death, and discharge, if applicable. Patients were followed up over a period of up to 7 days.

### Study Size

Due to the lack of published data on the frequency distribution of the relevant parameters in the hospitalized population, a formal sample size calculation was not feasible. We estimated that around 200 patients with smart devices and the corresponding data would be recruited based on the number of sites, the average amount of patients visiting the ED daily, and the percentage of people with a smart device [18,19,31]. This cohort size was considered representative in providing data with high external validity.

### Statistical Methods

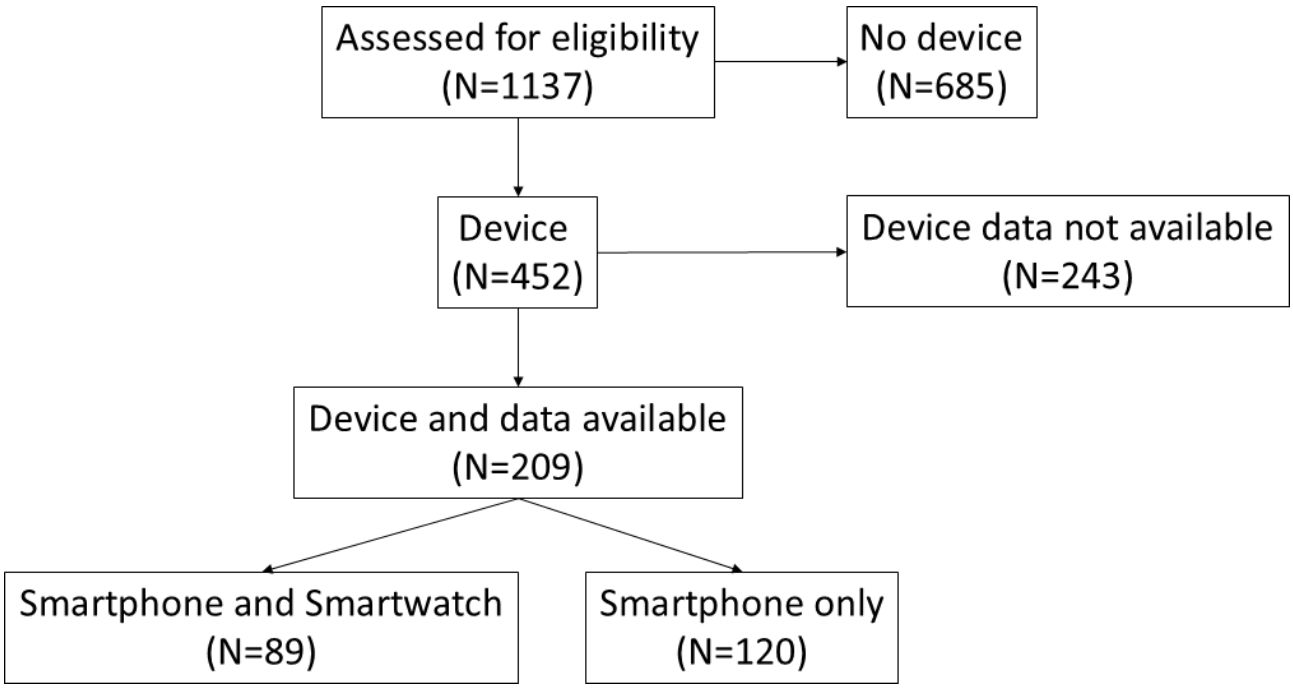
The primary outcome was feasibility, which was defined as having a smart device with heart rate or step count data recorded on at least twice between 30 days (baseline) before and on the day of admission, enabling trend analysis, and measured as the percentage of patients with an acute care visit. Patient characteristics were reported as absolute numbers per category or subcategory for the total patient population, discharged patients, and admitted patients. Variables were assessed for normal distribution. The heart rate and step count data were compared for their respective mean or median for the different time points. Heart rate changes were also assessed relative to baseline. Both were tested for significant interaction between time, disposition, and change in either heart rate or step count using a generalized linear model (GLM). To further assess the effect of time on change in heart rate and step count, a generalized linear mixed model (GLMM) with binomial distribution and random effects for country and hospital was conducted. Data were collected and safely stored using Castor EDC [32] and analyzed using SPSS (version 28, IBM Statistics). Statistical significance was defined at $\alpha$=.05.

## Results

### Feasibility

A total of 1137 patient cases were screened for this study (Figure 1). Of these, 685 were excluded for the following reasons: not having a smart device, not having their device with them during presentation to the hospital, not measuring health data with the smart device, and not having worn their device in the previous days due to illness. A total of 243 patients were excluded because of having insufficient amount of data, and 209 (18%) patients formed the core group for further analysis. Among these patients, 89 (43%) used a smartwatch (in combination with a smartphone) and 120 (57%) only used their smartphone (Multimedia Appendix 1). Heart rate data were recorded in 84 (40%) patients, and step count data were available in 207 (99%) patients.

**Figure 1.** Flowchart of the included patients.



### Characteristics of Subjects

We included a total of 209 patients who had either a smartwatch or smartphone that measured heart rate or step count in the week prior to their visit (Table 1). Distribution between male and female patients was equal. Most included patients were aged under 65 years, had attained higher education, and were confident in the usage of internet and applications. The overwhelming majority of the patients had low NEWS scores, low frailty scores, and normal gait. In terms of patient disposition, 75 (36%) patients were admitted and 126 (60%) were discharged, and for 8 (4%) patients, the outcome was unknown. Follow-up data showed that no patients died, 3 were admitted to the intensive care unit, and 122 were discharged within 7 days.

**Table .** Characteristics of included patients.

| Characteristics | All (n=209), n (%) | Admitted (n=75, 36%), n (%) | Discharged (n=126, 60%), n (%) |
|---|---|---|---|
| Gender | | | |
| Female | 109 (52) | 41 (55) | 65 (52) |
| Male | 100 (48) | 34 (45) | 61 (48) |
| Age group | | | |
| 18 - 30 y | 37 (18) | 14 (19) | 23 (18) |
| 31 - 50 y | 54 (26) | 14 (19) | 38 (30) |
| 51 - 65 y | 65 (31) | 25 (33) | 37 (29) |
| 65+ y | 53 (25) | 22 (29) | 28 (22) |
| Educational attainment | | | |
| Primary school | 15 (7) | 7 (9) | 7 (6) |
| Secondary school | 68 (33) | 24 (32) | 39 (31) |
| Higher education | 124 (59) | 42 (56) | 80 (63) |
| Unknown | 2 (1) | 2 (3) | 0 (0) |
| Confidence in internet usage | | | |
| Very | 125 (59) | 41 (55) | 80 (63) |
| Fairly | 65 (31) | 27 (36) | 35 (28) |
| Unsure | 10 (5) | 2 (3) | 7 (6) |
| Not very | 6 (3) | 3 (4) | 3 (2) |
| Not at all | 3 (1) | 2 (3) | 1 (1) |
| Confidence in app usage | | | |
| Very | 131 (63) | 43 (57) | 85 (67) |
| Fairly | 59 (28) | 24 (32) | 30 (24) |
| Unsure | 10 (5) | 4 (5) | 6 (5) |
| Not very | 6 (3) | 3 (4) | 3 (2) |
| Not at all | 2 (1) | 1 (1) | 1 (1) |
| Unknown | 1 (0) | 0 (0) | 1 (1) |
| National early warning score | | | |
| Low (0 - 2) | 178 (85) | 57 (76) | 115 (91) |
| High (3+) | 27 (13) | 18 (24) | 8 (6) |
| Unknown | 4 (2) | 0 (0) | 3 (2) |
| Clinical Frailty Scale | | | |
| 1-2 | 152 (73) | 44 (59) | 104 (83) |
| 3+ | 57 (27) | 31 (41) | 22 (17) |
| Normal gait | | | |
| Yes | 187 (89) | 63 (84) | 117 (93) |
| No | 17 (8) | 9 (12) | 7 (6) |
| Unknown | 5 (2) | 3 (4) | 2 (2) |

## Change in Heart Rate and Step Count

The mean resting heart rate increased over the 30 days preceding hospital attendance in all patient groups, including both admitted and discharged patients (Figure 2, Multimedia Appendix 2). Patients who were admitted had a higher mean resting heart rate than those who were discharged. The GLM showed a significant change in the mean resting heart rate over time ($P$<.001), but there was no significant interaction between time and patient disposition ($P$=.23). The GLMM correctly classified 85% (51/60) of cases, with higher accuracy for discharge (38/40, 95%) than for admission (13/20, 65%). Among the predictors,

only the mean resting heart rate on the day prior to presentation was significantly associated with patient disposition (β=–.083, SE=.038, *P*=.035). A higher resting heart rate on the day before an acute care visit was linked to a lower likelihood of admission (odds ratio 0.92 per beats per minute [bpm], 95% CI 0.85 - 0.99). Mean resting heart rate at 1 month (*P*=.66), 1 week (*P*=.23), on the day of presentation (*P*>.99), and during admission (*P*=.92) was not significantly associated with disposition. The percentage increase in the mean resting heart rate from baseline prior to hospital admittance was greater in admitted patients than in those who were discharged (Figure 3). In contrast, patients discharged after an acute care visit exhibited the highest percentile change in the mean resting heart rate.

**Figure 2.** Change in mean heart rate in days before an acute care visit. bpm: beats per minute; ED: emergency department.

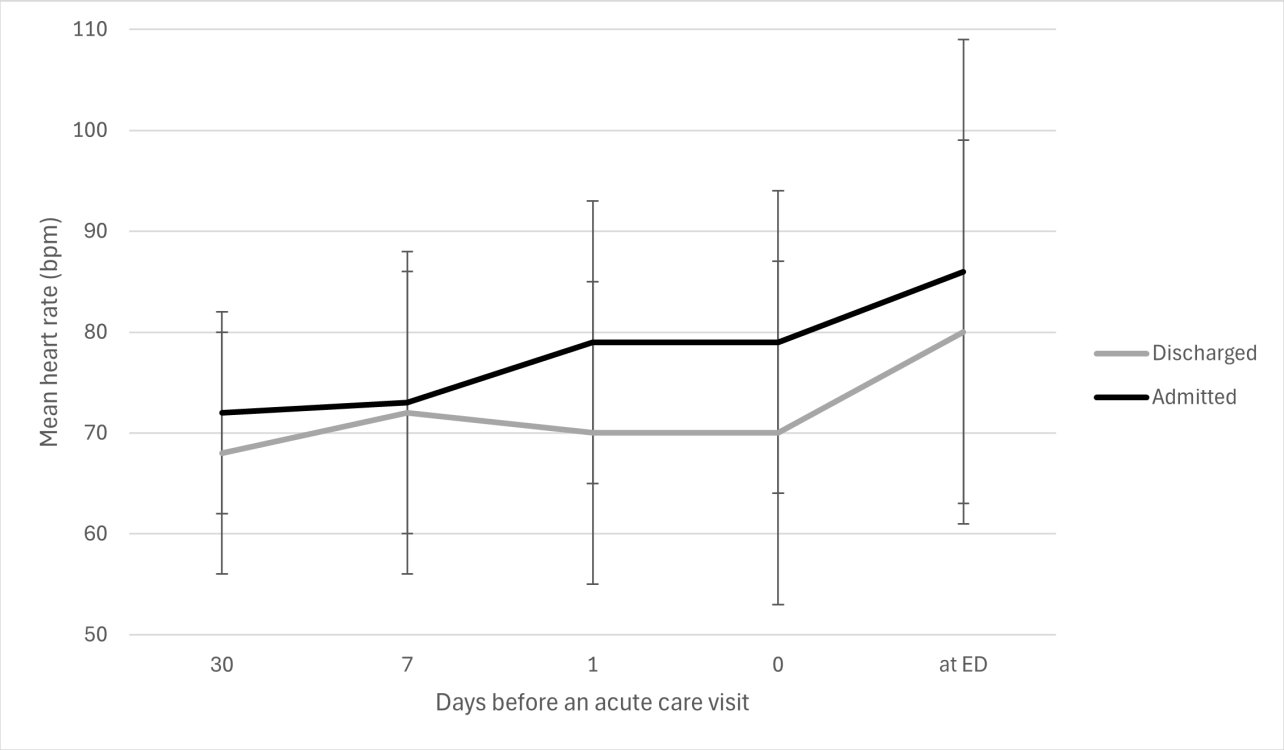
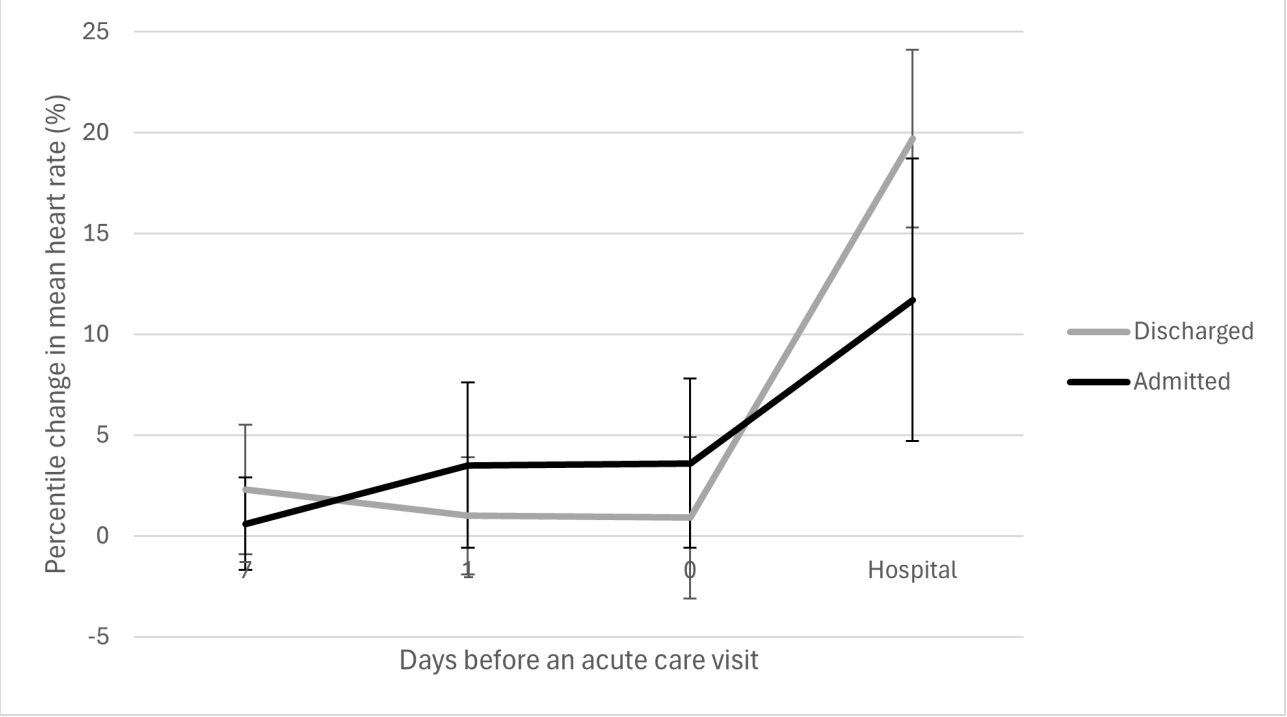


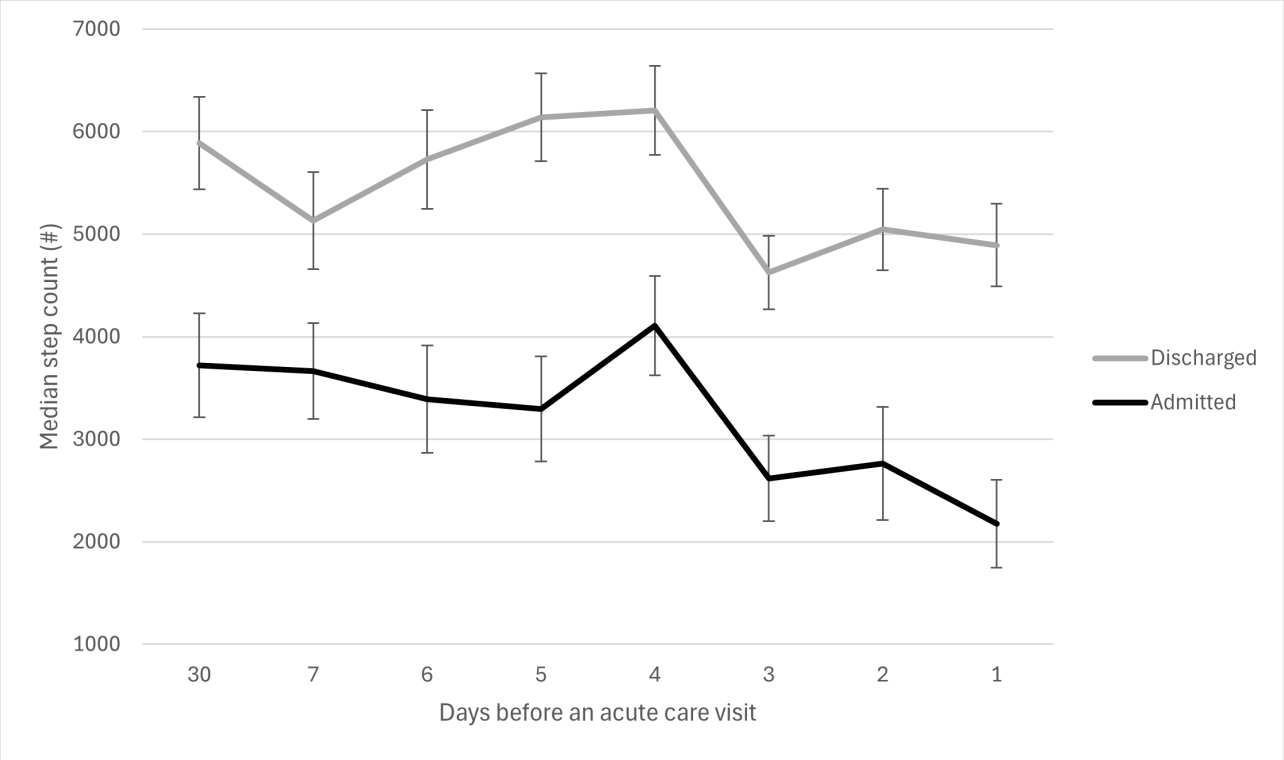

**Figure 3.** Percentile change in mean heart rate in days before an acute care visit.




The median step count decreased over the 30 days leading up to hospital presentation in both patient groups (Figure 4, Multimedia Appendix 3). The GLM showed a significant change in median step count over time (*P*<.001) but no significant

interaction between time and disposition (*P*=.10). The GLMM correctly classified 82% (112/136) of cases, with higher accuracy for discharge (83/88, 94%) than admission (29/48, 60%) patients. Among the predictors, only median step count on the day prior to presentation was significantly associated with disposition (β≈.0001, SE≈0.00009, *P*=.04). A lower median step count on the day before hospital presentation predicted an increased likelihood of admission (odds ratio=0.90 per 1000 steps, 95% CI 0.82 - 1.00). Median step counts at 30 (*P*=.43), 7 (*P*=.78), 6 (*P*=.74), 5 (*P*=.84), 4 (*P*=.22), 3 (*P*=.11), and 2 days (*P*=.18) prior to hospital presentation were not significantly associated with disposition.

**Figure 4.** Change in median step count in days before an acute care visit.




## Discussion

### Key Results

This study found that 40% of patients with an acute care visit had a smart device. In only 50% of the included patients, heart rate and step count data were available for trend analysis, resulting in data available in 20% of the total cohort. Within these patients, a significant increase in heart rate and a decrease in step count were observed several days before they sought emergency care.

### Interpretation

#### *Feasibility*

Feasibility was defined as having a smart device with heart rate or step count data recorded at least twice between 30 days (baseline) before and on the day of admission, enabling trend analysis, and measured as the percentage of patients with an acute care visit (ED, AMU, or SDEC). Device ownership, and thus feasibility, was lower in our study compared to reported daily use in the community (smartphone, 18% vs 90%; smartwatch, 8% vs 49%) [18,33-35]. We attribute this difference to two causes. First, the ED population is not representative of the general population. Patients who are at higher risk of requiring acute care are often older and from lower socioeconomic backgrounds [36,37], which contributes to lower device ownership [35]. Second, 40% had a smart device that measured heart rate and 99% had one that measured step count, leaving part of the potential value of health monitoring apps unused. This is more commonly seen in older patients [38]. Additionally, we excluded patients who did not have their smart device with them or had insufficient data (54%). Both were categorized as having no smart device. Nevertheless, in patients who used their smart device, the results demonstrated the potential clinical value of smart device recording of heart rate and step count data for assessing disease severity and rate of deterioration. These findings show that society is still a long way from converting care from analog to digital, but that the trend is slowly setting in.

#### *Change in Heart Rate and Step Count*

Changes in both step count and heart rate are associated with clinical deterioration [39,40]. In our study, we found that variations in these metrics on the day prior to an acute care visit significantly predicted patient disposition.

Interestingly, in our population, a greater increase in resting heart rate was associated with a lower likelihood of hospital admission. We hypothesize that this counterintuitive finding may be explained by the pattern of percentile change in resting heart rate (Figure 3). Discharged patients tended to exhibit a brief, sharp increase, whereas admitted patients showed a more gradual and prolonged rise. This difference may reflect variations in autonomic regulation: healthier individuals—those

ultimately discharged—could have greater heart rate variability (HRV), allowing for more pronounced fluctuations. Although HRV was not directly measured in our study, literature suggests that higher HRV is associated with better health status and resilience to stressors, potentially explaining the observed spike in heart rate in discharged patients [41,42]. Though, it is important to note that the group of discharged patients maintained a lower overall mean resting heart rate compared to those who were admitted.

Other studies have similarly demonstrated that changes in heart rate are linked to illness severity and that failure to normalize heart rate by the time of disposition is associated with worse outcomes [1,41-45]. Furthermore, the mean resting heart rate increased in both groups, although it remained within the normal range defined by early warning scores such as NEWS.

Notably, heart rate measured by the patient's device on the day of the acute care visit and in the hospital during triage (ED, AMU, or SDEC) showed an increase in both discharged (from 75 to 82 bpm) and admitted (from 79 to 88 bpm) patients. This increase begins on day −1 and appears predictive of admission. The rise may reflect acute physiological stress due to illness, emotional stress associated with hospital visits, or differences in measurement context. While discrepancies between consumer-grade devices and hospital equipment could be a factor, prior validation studies suggest that smartwatches and activity trackers provide sufficiently accurate measurements [46,47], making this explanation less likely.

The overall difference between admitted and discharged patients highlights a substantial disparity in general physical fitness between the 2 groups. Compared to discharged patients, those who were admitted had a significantly lower overall median step count, with a further decline beginning 2 days prior to the acute care visit. This pattern may enable the identification of patients at risk of admission as early as a day before presentation. These findings support previous findings that physical activity, including step count, tends to decline as patients become more acutely ill [3,48,49].

These findings underscore the importance of using individualized baselines and relative changes as reference points, highlighting the potential of personalized medicine.

### *Study Population*

Our final study population was reflective of the general acute care population for both age and gender but reported a higher level of education [50]. Higher education usually corresponds with a higher socioeconomic class, and a higher socioeconomic class has a lowered risk of acute care visits and poor health outcomes [36,37]. Furthermore, our population was confident in the usage of both the internet and applications, indicating strong digital health literacy [51]. This is most likely because a higher education is associated with a higher digital health literacy [52]. The included patients had mostly low NEWS, CFS, and normal gait. This indicates that our participants were relatively healthy compared to the average acute care population. This is supported by a recent European study showing that 40% of older people using the emergency care had CFS 5+ [53]. Furthermore, many patients these days visit the ED for acute complaints of chronic diseases, demonstrating the increasing fragility of the ED population [54]. Based on our findings, it seems the patients who would potentially benefit the most from digital health monitoring do not use it, as opposed to the relatively healthy patients who already use it. In summary, we included a group of patients from a high socioeconomic class in relatively good health, which is not consistent with the general population of acute patients.

## Strengths and Limitations

The major strength is the generalizability of this point prevalence study into the use of smart devices in Northwestern Europe, including several countries, regions, and hospital settings. Additionally, our study highlights the clinical value of patient-held sensors and understanding patients' own reference value. This could be used to track a patient's health remotely and proactively (ie, monitoring individuals before they have an acute illness and become patients) by both patients and health care providers.

Our study had several limitations. First, by only including patients who had a smart device and sufficient data, we created a selection bias. We selected patients with higher education and most likely a higher socioeconomic status, representing a healthier population. However, we expect the underlying pathophysiology to be unaffected by this selection. Second, because of the flash mob research design and because we wanted to minimize the workload for the participating centers, we chose to collect only a limited number of parameters. Resting heart rate between 6 and 2 days before an acute care visit, reason for attendance, medication use, medical history, psychosocial aspects, and ethnicity were not considered even though these parameters could have provided more insight into which groups of patients are being admitted. Finally, the sample size was relatively small, making the subgroup analysis (eg, device type) not feasible. The small sample size and selection bias toward individuals with less compromised physiology may lead to the underestimation of the changes before and on the day of contact with the emergency services.

## Future Perspectives

Proactive monitoring of health data in populations at high risk of acute illness could enable earlier identification of deteriorating health. Monitoring on a daily basis may help detect subtle changes in vital signs or physical activity, allowing for timely medical assessment in primary or ambulatory care settings and potentially preventing a hospital visit. This strategy is already recommended for atrial fibrillation detection by the European Society of Cardiology guidelines 2024 [55]. This benefit is potentially greatest for frail patients, but they are less likely to use a smart device to monitor their health. In addition, changes in vital signs may be more subtle in frail patients, making it more difficult to detect changes. In fitter patients, changes in heart rate and step count may be steeper, but as shown in our study population, the theoretical benefit in a nonselected population is limited. Overall, patients might be encouraged to bring their smart devices with them to an acute care visit, and clinicians might consider asking patients to look at these smart devices, as information on recent heart rate and step counts may be helpful, which can be used as a valid

measurement of resting heart rate, as shown by several studies [46,47]. Additionally, both users and manufacturers of smart devices can help health care providers by expanding their reach. Currently, users tend to be younger, more educated, and digitally literate patients, while devices could show clinically valuable trends for all. Both patients and professionals should be wary of misdiagnosis or overdiagnosis. This study provides a foundation for further research to help patients and clinicians pick up early deterioration in health, which might attract interest from tech companies and mobile application developers. Furthermore, a targeted information campaign could be considered, which should be aimed at high-risk patient groups.

## Conclusions

Our study showed that it is feasible to use a patient's own smart device to measure vital signs in the days preceding an acute care visit. We found a significant change in heart rate and step count prior to presentation to hospital, where disposition can be predicted using a patient's own smart device data. These smart devices are mostly used by younger and healthier patients with higher educational attainment. The use of a patient's own smart devices for health monitoring in high-risk patient groups is very limited.

## Acknowledgments

The study was initiated and conducted by a collaboration of the Safer@Home consortium and the Research Consortium Acute Medicine (ORCA). We would like to thank all participating patients and health care professionals. We especially thank the local coordinators of all participating centers listed in Multimedia Appendix 4.

## Data Availability

The anonymized dataset is available on reasonable request after approval of the corresponding author.

## Authors' Contributions

Conceptualization: JA (equal), CPS (equal)
Data Curation: JGAD
Formal Analysis: JGAD (lead), AAMH (supporting), CSP (supporting), JA (supporting)
Methodology: JA (equal), JK (supporting), EFC (supporting), HRH (supporting), MB (supporting), CHN (supporting), PWBN (supporting), and CPS (equal)
Project Administration: JGAD (lead) and JA (supporting)
Supervision: JA (equal), CPS (equal), MB (equal), and CHN (equal)
Visualization: JGAD (lead), AAMH (supporting), CSP (supporting), JA (supporting)
Writing – Original Draft: JGAD (equal), AAMH (equal), CSP (equal), JA (equal)
Writing – Review & Editing: JGAD (lead), AAMH (supporting), JK (supporting), EFC (supporting), HRH (supporting), MB (supporting), CHN (supporting) PWBN (supporting), CPS (supporting), and JA (supporting)

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of devices of included patients.
[XLSX File, 13 KB - cardio_v9i1e76218_app1.xlsx ]

Multimedia Appendix 2
Change in heart rate in patients.
[XLSX File, 13 KB - cardio_v9i1e76218_app2.xlsx ]

Multimedia Appendix 3
Change in step counts in patients.
[XLSX File, 13 KB - cardio_v9i1e76218_app3.xlsx ]

Multimedia Appendix 4
Contributing author list.
[DOCX File, 16 KB - cardio_v9i1e76218_app4.docx ]

## References

1. Brekke IJ, Puntervoll LH, Pedersen PB, Kellett J, Brabrand M. The value of vital sign trends in predicting and monitoring clinical deterioration: a systematic review. PLoS ONE 2019;14(1):e0210875. [doi: 10.1371/journal.pone.0210875] [Medline: 30645637]
2. Zhang D, Shen X, Qi X. Resting heart rate and all-cause and cardiovascular mortality in the general population: a meta-analysis. CMAJ 2016 Feb 16;188(3):E53-E63. [doi: 10.1503/cmaj.150535] [Medline: 26598376]
3. Rice H, Hill K, Fowler R, Watson C, Waterer G, Harrold M. Reduced step count and clinical frailty in hospitalized adults with community-acquired pneumonia. Respir Care 2020 Apr;65(4):455-463. [doi: 10.4187/respcare.06992] [Medline: 31575707]
4. Helmer P, Hottenrott S, Rodemers P, et al. Accuracy and systematic biases of heart rate measurements by consumer-grade fitness trackers in postoperative patients: prospective clinical trial. J Med Internet Res 2022 Dec 30;24(12):e42359. [doi: 10.2196/42359] [Medline: 36583938]
5. van der Kluit MJ, Dijkstra GJ, de Rooij SE. The decision-making process for unplanned admission to hospital unveiled in hospitalised older adults: a qualitative study. BMC Geriatr 2018 Dec 22;18(1):318. [doi: 10.1186/s12877-018-1013-y] [Medline: 30577791]
6. Emmanuel A, Ismail A, Kellett J. Assessing the need for hospital admission by the Cape Triage discriminator presentations and the simple clinical score. Emerg Med J 2010 Nov;27(11):852-855. [doi: 10.1136/emj.2009.086256] [Medline: 20513735]
7. Subbe CP, Jishi F, Hibbs RAB. The simple clinical score: a tool for benchmarking of emergency admissions in acute internal medicine. Clin Med (Lond) 2010 Aug;10(4):352-357. [doi: 10.7861/clinmedicine.10-4-352] [Medline: 20849009]
8. Kellett J, Emmanuel A, Deane B. Who will be sicker in the morning? Changes in the simple clinical score the day after admission and the subsequent outcomes of acutely ill unselected medical patients. Eur J Intern Med 2011 Aug;22(4):375-381. [doi: 10.1016/j.ejim.2011.03.005] [Medline: 21767755]
9. Bleyer AJ, Vidya S, Russell GB, et al. Longitudinal analysis of one million vital signs in patients in an academic medical center. Resuscitation 2011 Nov;82(11):1387-1392. [doi: 10.1016/j.resuscitation.2011.06.033] [Medline: 21756971]
10. Jones M. NEWSDIG: the National Early Warning Score development and implementation group. Clin Med (Lond) 2012 Dec;12(6):501-503. [doi: 10.7861/clinmedicine.12-6-501] [Medline: 23342400]
11. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax 2003 May;58(5):377-382. [doi: 10.1136/thorax.58.5.377] [Medline: 12728155]
12. Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for upper-gastrointestinal haemorrhage. Lancet 2000 Oct 14;356(9238):1318-1321. [doi: 10.1016/S0140-6736(00)02816-6] [Medline: 11073021]
13. Hosseini M, Baikpour M, Yousefifard M, et al. Blood pressure percentiles by age and body mass index for adults. EXCLI J 2015;14:465-477. [doi: 10.17179/excli2014-635] [Medline: 26417366]
14. Chester JG, Rudolph JL. Vital signs in older patients: age-related changes. J Am Med Dir Assoc 2011 Jun;12(5):337-343. [doi: 10.1016/j.jamda.2010.04.009] [Medline: 21450180]
15. Berge HM, Isern CB, Berge E. Blood pressure and hypertension in athletes: a systematic review. Br J Sports Med 2015 Jun;49(11):716-723. [doi: 10.1136/bjsports-2014-093976] [Medline: 25631543]
16. Lundstrom CJ, Foreman NA, Biltz G. Practices and applications of heart rate variability monitoring in endurance athletes. Int J Sports Med 2023 Jan;44(1):9-19. [doi: 10.1055/a-1864-9726] [Medline: 35853460]
17. Grant S, Crimmons K. Limitations of track and trigger systems and the National Early Warning Score. Part 2: sensitivity versus specificity. Br J Nurs 2018 Jun 28;27(12):705-710. [doi: 10.12968/bjon.2018.27.12.705] [Medline: 29953281]
18. Laricchia F. Smartphone ownership in the US by age 2023. : Statista; 2024 URL: www.statista.com/statistics/489255/percentage-of-us-smartphone-owners-by-age-group [accessed 2024-11-19]
19. Laricchia F. Smartphone ownership penetration in the united kingdom (UK) in 2012-2024, by age. : Statista; 2024 URL: https://www.statista.com/statistics/271851/smartphone-owners-in-the-united-kingdom-uk-by-age [accessed 2025-02-26]
20. Bin KJ, De Pretto LR, Sanchez FB, Battistella LR. Digital platform to continuously monitor patients using a smartwatch: preliminary report. JMIR Form Res 2022 Sep 15;6(9):e40468. [doi: 10.2196/40468] [Medline: 36107471]
21. Wettstein R, Sedaghat-Hamedani F, Heinze O, et al. A remote patient monitoring system with feedback mechanisms using a smartwatch: concept, implementation, and evaluation based on the activeDCM randomized controlled trial. JMIR Mhealth Uhealth 2024 Nov 22;12:e58441. [doi: 10.2196/58441] [Medline: 39365164]
22. Buda KG, Sengupta J. Smartwatches and atrial fibrillation: what works and what needs improvement? : American College of Cardiology; 2024 URL: https://www.acc.org/latest-in-cardiology/articles/2024/05/29/16/56/smartwatches-and-atrial-fibrillation [accessed 2025-08-22]
23. Li X, Dunn J, Salins D, et al. Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. PLoS Biol 2017 Jan;15(1):e2001402. [doi: 10.1371/journal.pbio.2001402] [Medline: 28081144]
24. Mishra T, Wang M, Metwally AA, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. Nat Biomed Eng 2020 Dec;4(12):1208-1220. [doi: 10.1038/s41551-020-00640-6] [Medline: 33208926]
25. Gadaleta M, Radin JM, Baca-Motes K, et al. Passive detection of COVID-19 with wearable sensors and explainable machine learning algorithms. NPJ Digit Med 2021 Dec 8;4(1):166. [doi: 10.1038/s41746-021-00533-1] [Medline: 34880366]

26. Ding EY, Tran KV, Lessard D, et al. Accuracy, usability, and adherence of smartwatches for atrial fibrillation detection in older adults after stroke: randomized controlled trial. JMIR Cardio 2023 Nov 28;7:e45137. [doi: 10.2196/45137] [Medline: 38015598]
27. Köhler C, Bartschke A, Fürstenau D, Schaaf T, Salgado-Baez E. The value of smartwatches in the health care sector for monitoring, nudging, and predicting: viewpoint on 25 years of research. J Med Internet Res 2024 Oct 25;26:e58936. [doi: 10.2196/58936] [Medline: 39356287]
28. Alsma J, van Saase JLCM, Nanayakkara PWB, et al. The power of flash mob research: conducting a nationwide observational clinical study on capillary refill time in a single day. Chest 2017 May;151(5):1106-1113. [doi: 10.1016/j.chest.2016.11.035] [Medline: 27940191]
29. Barrington J, Subbe C, Kellett J, et al. Feasibility of using resting heart rate and step counts from patient-held sensors during clinical assessment of medical emergencies (FUSE): protocol for prospective observational study in European hospitals. JMIR Res Protoc 2025 Apr 28;14:e55975. [doi: 10.2196/55975] [Medline: 40293791]
30. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005 Aug 30;173(5):489-495. [doi: 10.1503/cmaj.050051] [Medline: 16129869]
31. Gaakeer MI, Veugelers R, van Lieshout JM, Patka P, Huijsman R. The emergency department landscape in The Netherlands: an exploration of characteristics and hypothesized relationships. Int J Emerg Med 2018 Sep 5;11(1):35. [doi: 10.1186/s12245-018-0196-5] [Medline: 31179931]
32. Castor electronic data capture. Castor EDC. URL: https://www.castoredc.com/electronic-data-capture-system/ [accessed 2025-11-14]
33. Laricchia F. Smartphone ownership by age 2012-2023. : Statista; 2024 URL: https://www.statista.com/statistics/271851/smartphone-owners-in-the-united-kingdom-uk-by-age [accessed 2024-11-19]
34. Bashir U. Wearable devices ownership in selected countries as of march 2024. : Statista; 2024 URL: https://www.statista.com/forecasts/1101101/wearable-devices-ownership-in-selected-countries [accessed 2025-02-05]
35. Kim E, Torous J, Horng S, et al. Mobile device ownership among emergency department patients. Int J Med Inform 2019 Jun;126:114-117. [doi: 10.1016/j.ijmedinf.2019.03.020] [Medline: 31029252]
36. Jonassaint CR, Beach MC, Haythornthwaite JA, et al. The association between educational attainment and patterns of emergency department utilization among adults with sickle cell disease. Int J Behav Med 2016 Jun;23(3):300-309. [doi: 10.1007/s12529-016-9538-y] [Medline: 26818355]
37. Lee SY, Lee SR, Choi EK, Han KD, Oh S, Lip GYH. Impact of socioeconomic status on emergency department visits in patients with atrial fibrillation: a nationwide population-based cohort study. J Am Heart Assoc 2022 Dec 20;11(24):e027192. [doi: 10.1161/JAHA.122.027192] [Medline: 36515229]
38. Malani P, Kullgren J, Solway E, Lee P, Aikens J, Richardson C, et al. Mobile health app use among older adults. : Institute for Healthcare Policy & Innovation, University of Michigan; 2022 URL: https://www.healthyagingpoll.org/reports-more/report/mobile-health-app-use-among-older-adults [accessed 2025-02-05]
39. Moy ML, Teylan M, Weston NA, Gagnon DR, Garshick E. Daily step count predicts acute exacerbations in a US cohort with COPD. PLoS ONE 2013;8(4):e60400. [doi: 10.1371/journal.pone.0060400] [Medline: 23593211]
40. Mayampurath A, Jani P, Dai Y, Gibbons R, Edelson D, Churpek MM. A vital sign-based model to predict clinical deterioration in hospitalized children. Pediatr Crit Care Med 2020 Sep;21(9):820-826. [doi: 10.1097/PCC.0000000000002414] [Medline: 32511200]
41. Joyce D, Barrett M. State of the science: heart rate variability in health and disease. BMJ Support Palliat Care 2019 Sep;9(3):274-276. [doi: 10.1136/bmjspcare-2018-001588] [Medline: 30446491]
42. Martinez GJ, Grover T, Mattingly SM, et al. Alignment between heart rate variability from fitness trackers and perceived stress: perspectives from a large-scale in situ longitudinal study of information workers. JMIR Hum Factors 2022 Aug 4;9(3):e33754. [doi: 10.2196/33754] [Medline: 35925662]
43. Singh B, Cadenas-Sanchez C, da Costa BGG, et al. Comparison of objectively measured and estimated cardiorespiratory fitness to predict all-cause and cardiovascular disease mortality in adults: a systematic review and meta-analysis of 42 studies representing 35 cohorts and 3.8 million observations. J Sport Health Sci 2024 Sep 11;14:100986. [doi: 10.1016/j.jshs.2024.100986] [Medline: 39271056]
44. Griffin F, Ekblom-Bak E, Arvidsson D, Paulsson S, Börjesson M. Maintaining or increasing cardiorespiratory fitness is associated with reduced hospital admission rate. Eur J Prev Cardiol 2024 Mar 4;31(4):436-444. [doi: 10.1093/eurjpc/zwad367] [Medline: 38019572]
45. Natarajan A, Su HW, Heneghan C. Assessment of physiological signs associated with COVID-19 measured using wearable devices. NPJ Digit Med 2020 Nov 30;3(1):156. [doi: 10.1038/s41746-020-00363-7] [Medline: 33299095]
46. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 8;8(9):e18694. [doi: 10.2196/18694] [Medline: 32897239]
47. Montalvo S, Martinez A, Arias S, et al. Commercial smart watches and heart rate monitors: a concurrent validity analysis. J Strength Cond Res 2023 Sep 1;37(9):1802-1808. [doi: 10.1519/JSC.0000000000004482] [Medline: 36862131]

48. Brouwer CG, Bartelet BM, Douma JAJ, et al. Machine learning-based prediction of clinical outcomes in patients with cancer receiving systemic treatment using step count data measured with smartphones. JCO Clin Cancer Inform 2025 Jul;9:e2500023. [doi: 10.1200/CCI-25-00023] [Medline: 40587817]
49. Zisman-Ilani Y, Fasing K, Weiner M, Rubin DJ. Exercise capacity is associated with hospital readmission among patients with diabetes. BMJ Open Diabetes Res Care 2020 Oct;8(1):e001771. [doi: 10.1136/bmjdrc-2020-001771] [Medline: 33020136]
50. Candel BG, Dap S, Raven W, et al. Sex differences in clinical presentation and risk stratification in the emergency department: an observational multicenter cohort study. Eur J Intern Med 2022 Jan;95:74-79. [doi: 10.1016/j.ejim.2021.09.001] [Medline: 34521584]
51. van Kessel R, Wong BLH, Clemens T, Brand H. Digital health literacy as a super determinant of health: more than simply the sum of its parts. Internet Interv 2022 Mar;27:100500. [doi: 10.1016/j.invent.2022.100500] [Medline: 35242586]
52. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
53. European Taskforce on Geriatric Emergency Medicine (ETGEM) collaborators. Prevalence of frailty in European emergency departments (FEED): an international flash mob study. Eur Geriatr Med 2024 Apr;15(2):463-470. [doi: 10.1007/s41999-023-00926-3] [Medline: 38340282]
54. NIV. Toekomstbestendige acute zorg reikt verder dan de spoedeisende hulp [Article in Dutch]. 2023 URL: https://www.internisten.nl/wp-content/uploads/2023/02/2023-02-6-3a2-NIV-visie-toekomstbestendige-acute-zorg.pdf [accessed 2025-11-13]
55. Van Gelder IC, Rienstra M, Bunting KV, et al. 2024 ESC Guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2024 Sep 29;45(36):3314-3414. [doi: 10.1093/eurheartj/ehae176] [Medline: 39210723]


## Abbreviations

**AMU:** acute medical unit
**bpm:** beats per minute
**CFS:** Clinical Frailty Scale
**ED:** emergency department
**GLM:** generalized linear model
**GLMM:** generalized linear mixed model
**HRV:** heart rate variability
**NEWS:** National Early Warning Score
**SDEC:** same day emergency care


*Edited by J Rivers; submitted 18.04.25; peer-reviewed by M Snyder, V Kan; revised version received 19.09.25; accepted 05.10.25; published 15.12.25.*

*Please cite as:*
*den Duijn JGA, Hajjaj AAM, Kellett J, Frischknecht Christensen E, Haak HR, Brabrand M, Nickel CH, Nanayakkara PWB, Subbe CP, Alsma J, Safer@Home, Research Consoritum Acute Medicine (ORCA)*
*Using Patient-Held Devices to Measure Variations in Resting Heart Rate and Step Count Prior to Presentation With an Acute Illness: International, Multicenter Flash Mob Feasibility Study*
*JMIR Cardio 2025;9:e76218*
*URL: https://cardio.jmir.org/2025/1/e76218*
*doi:10.2196/76218*

# Web-Based Platform for the Chilean Cardiac Surgery Registry: Algorithm Development and Validation Study

Sergio Guinez-Molinos[1*], PhD; Enrique Seguel[2,3*], MD; Jaime Gonzalez[1*], MSc; Benjamin Castillo[4*], PhD

[1]Simulation and Biomedical Informatics Laboratory, Faculty of Medicine, Universidad de Talca, San Miguel 3748, PO Box 721, Talca, Chile

[2]Cardiovascular Centre, Guillermo Grant Benavente Hospital, Concepción, Chile

[3]Department of Surgery, Faculty of Medicine, University of Concepción, Concepción, Chile

[4]Centro de Investigación de Estudios Avanzados del Maule (CIEAM), Vicerrectoría de Investigación y Postgrado, Universidad Católica del Maule, Talca, Chile

[*]all authors contributed equally

**Corresponding Author:**
Sergio Guinez-Molinos, PhD
Simulation and Biomedical Informatics Laboratory, Faculty of Medicine, Universidad de Talca, San Miguel 3748, PO Box 721, Talca, Chile

## Abstract

**Background:** Cardiac surgeries in Chile lack a national registry for systematic data collection and analysis, limiting insights into procedural outcomes and patient demographics. In response to this gap, we developed a web-based platform to support the documentation of high-complexity cardiac surgeries.

**Objective:** This study aimed to design, develop, and implement a cardiac surgery data collection and analysis platform that conforms to international standards to support clinical decision-making and research initiatives.

**Methods:** A web-based platform was developed using the model-view-controller architecture, incorporating input from health care professionals and based on the fourth European Association for Cardio-Thoracic Surgery adult cardiac surgical database report. The platform captures more than 160 clinical variables across 15 categories, spanning preoperative, intraoperative, and postoperative stages.

**Results:** The most significant outcome of this study is the development of the first online platform for documenting cardiac surgeries in Chile. Since its implementation in 2014, the platform has documented more than 4800 cardiac surgeries, establishing it as the largest database for a single institution in Latin America. The platform offers real-time access to data, supports planning and resource allocation, and enables the systematic evaluation of clinical outcomes. Integrating the European System for Cardiac Operative Risk Evaluation II risk model enables a standardized assessment of mortality risk.

**Conclusions:** The platform contributes to the collection of cardiac surgery data in Chile, enabling evidence-based clinical decision-making and informed public health planning. It has documented cardiac surgeries for 10 years and has become the official registry tool for cardiac surgeries. By 2026, its application will be extended to 2 more centers, with the expectation that it will soon become the national database of cardiac surgeries. Future developments should improve scalability, interoperability, and data analysis to establish a national registry and further align Chilean cardiac surgery practices with international standards.

***(JMIR Cardio 2025;9:e70147)*** doi:10.2196/70147

**KEYWORDS**

cardiac surgery; health informatics; clinical registry; EuroSCORE II; Chile

## Introduction

Integrating health information systems into clinical practice has transformed cardiac surgery, establishing a basis for evidence-based decision-making [1]. Electronic medical records, easier access to clinical data, and statistical analysis tools have improved the approaches to treating and managing various cardiac conditions [2].

International scientific societies have recognized the importance of standardized data collection and have developed electronic databases to collect and analyze cardiac surgery outcomes. The Society of Thoracic Surgeons database in the United States has become the largest cardiac surgery registry in the world [3]. Other initiatives include the databases created by the Spanish Society of Thoracic and Cardiovascular Surgery [4], the German Society of Thoracic and Cardiovascular Surgery [5], the British Society of Cardio-Thoracic Surgeons [6], the Australian and New Zealand Society of Cardio-Thoracic Surgeons [7], and the Japanese Society of Cardiovascular Surgery [8]. In Europe, the European Association for Cardio-Thoracic Surgery (EACTS) [9,10], with the support of the Society of Thoracic Surgeons,

has also made significant contributions to the worldwide network of cardiac surgery registries. These databases highlight the importance of systematic registries in understanding the epidemiology and outcomes of cardiac surgery [11].

In Chile, there is no national epidemiological registry or health information system that systematically gathers data on cardiac surgery and related procedures [11]. As a result, the actual number of procedures performed, their short- and long-term outcomes, and the demographic and clinical profiles of patients remain largely unknown. This lack of structured data hampers the objective evaluation of health interventions, limits follow-up efforts, and restricts the development of evidence-based public health policies. Consequently, decision-making in this field often becomes reactive, lacking the necessary information to assess impact or allocate resources efficiently and accurately.

To address these challenges, the first web-based cardiac surgery registry platform in Chile has been developed and implemented. Specifically designed to document highly complex cardiac procedures, this initiative emerged from the joint efforts of the Cardiac Surgery Service at Guillermo Grant Benavente Hospital, the Department of Surgery at the University of Concepción, and the Center for Simulation and Biomedical Informatics at the Faculty of Medicine of the University of Talca. The primary aim of this platform is to bridge the current data gap and establish a robust foundation for enhancing surgical outcomes in cardiac care nationwide.

Currently, Guillermo Grant Benavente Hospital is the only center fully integrated into this electronic registry. In parallel, formal discussions with the Ministry of Health are underway, with a long-term vision of scaling this initiative into a nationwide platform that can serve as the cornerstone for monitoring quality, benchmarking outcomes, and guiding clinical decision-making in Chilean cardiac surgery.

## Methods

A standard software engineering methodology was used to develop a platform for recording cardiac surgery and procedure data [12]. Close collaboration with cardiology experts was essential from the outset, with nurses and physicians playing a key role in the iterative requirements-gathering process. The platform's scope, modules, and sections were defined through a series of meetings at the Cardiac Surgery Registry (Table 1). This collaborative approach, involving different clinical and IT profiles such as cardiac surgeons, cardiologists, perfusionists, nurses, and biomedical informatics specialists, resulted in a detailed technical document that served as the basis for the software's development.

**Table .** Summary of the requirements survey, with the main functional and nonfunctional requirements to be covered by the platform.

| Modules | Requirements | Detailed requirements |
|---|---|---|
| Platform administration | Functional | • The system must implement user authorization and authentication. User accounts must have profiles within the platform according to their roles in the medical process registry. Each role should be associated with specific privileges or functionalities.<br>• The platform should provide functionalities to create, update, and delete user accounts, manage hospital lists, and refer health care services.<br>• Users can have different roles: administrator, professional, or data logger. |
| Cardiac surgery registry | Functional | • The system should be capable of storing data related to complex cardiac surgeries and procedures. The available input fields should be based on the EACTS[a] dataset modified to suit local requirements.<br>• The system should also automatically calculate indicators such as EuroSCORE II[b] to support decision-making processes.<br>• Multilanguage (English and Spanish). |
| Patients, institutions, and inventory | Functional | • The system should securely store patient data, ensuring anonymity, and avoiding records.<br>• It should also store data related to personnel, institutions, and equipment used during interventions and patient monitoring.<br>• The recorded data must be interoperable, allowing it to be shared with other platforms in the ecosystem. |
| Reporting and data recovery | Functional | • The system must allow the export of data related to the procedures.<br>• It must be able to generate reports and visualizations of the data recorded and the use of the platform.<br>• Data retrieval and validation must be guaranteed. |
| All modules | Nonfunctional | • Ease of use and simplicity are essential, especially for data entry associated with surgeries or medical procedures and visualization.<br>• Data security must be ensured with strong authorization, authentication, and privacy measures.<br>• Fail-safe mechanisms must be implemented to ensure the accuracy of data capture and prevent data loss or inconsistencies.<br>• The platform must support scalability and data independence, allowing the addition of new data categories without affecting existing data or data capture processes.<br>• It must also facilitate seamless integration with other platforms. |

[a]EACTS: European Association for Cardio-Thoracic Surgery.

[b]EuroSCORE II: European System for Cardiac Operative Risk Evaluation.

The analysis of clinical requirements highlighted 2 essential elements for the platform's design and development. First, the EACTS guidelines should be followed [13], considering the clinical variables recorded in that database and adapting the registry to local needs. Second, the platform should incorporate automatic calculation of the European System for Cardiac

Operative Risk Evaluation II (EuroSCORE II) mortality risk indicator to support clinical decision-making processes [10,14].

## Functional and Nonfunctional Requirements

The cardiac surgery registry is designed for health care professionals to record and query data related to medical procedures. Administrators manage the platform's maintenance, including user privileges and master tables. Both cardiac surgeons and nurses who are part of the clinical procedural team can enter data related to complex surgeries and procedures. Surgeons are associated with specific guidelines and can review the corresponding records from anywhere with an internet connection.

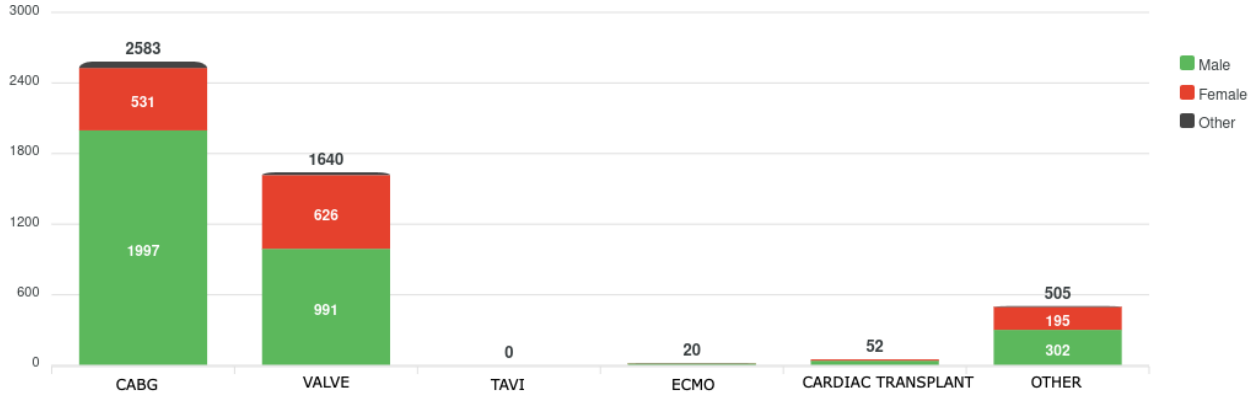
A fundamental aspect is the security of sensitive data. A data security layer is implemented that guarantees, in accordance with national data protection legislation, the appropriate authorization, authentication, and privacy of records.

The platform features multilevel user authentication, encrypted connections (secure sockets layer and transport layer security) [15], anonymization of patient identifiers, and storage on secure servers with regular backups. Access is based on roles (surgeons, nurses, and administrators) and is designed to restrict unauthorized access to sensitive information.

## Cardiac Surgery Clinical Sections

The types of cardiac surgeries included in the registry are as follows: coronary artery bypass grafting (CABG), valve surgeries (aortic, mitral, and tricuspid), and combined CABG and valve procedures (Figure 1). In addition, extracorporeal membrane oxygenation and heart transplants are being implemented on a pilot basis.

**Figure 1.** Chart illustrating the distribution of procedures by type according to gender. CABG: coronary artery bypass grafting; ECMO: extracorporeal membrane oxygenation; TAVI: transcatheter aortic valve implantation.




According to the surgeons' and their teams' requirements, clinical sections were established for data collection on cardiovascular surgeries. Key categories included patient demographics, cardiovascular history, previous interventions, preoperative risk factors, hemodynamics, and immediate status before surgery. Detailed information on the procedure, echocardiogram findings, and myocardial protective measures was also documented, ensuring a thorough understanding of the surgical and intraoperative contexts. Postoperative complications, discharge outcomes, and long-term patient follow-up were systematically tracked to assess recovery and survival rates. All the above are based on the EACTS structure [10], from which we derived our clinical sections (Table 2).

**Table .** The platform will represent the main clinical sections.

| Clinical section | Description |
|---|---|
| Hospitalization details | Captures patient identification, admission specifics, health care service, and the urgency of intervention |
| Cardiovascular history | Documents heart-related conditions such as angina, myocardial infarctions, and congestive heart failure |
| Previous interventions | Records prior surgeries, angioplasties, and their dates |
| Preoperative risk factors | Assesses risks, including weight, smoking history, and preexisting medical conditions |
| Preoperative hemodynamics and catheterization | Includes diagnostic metrics such as coronary vessel status and ejection fraction |
| Preoperative status and support | Notes presurgery interventions such as IV medications and mechanical support |
| Operation details | Provides surgical specifics, including type, urgency, and personnel involved |
| Coronary surgery | If a coronary surgery was performed, it is registered here |
| Valve surgery | Stenosis, insufficiency, explant type, and other data are registered for valve surgery (aortic, mitral, tricuspid, and pulmonary) |
| Echocardiogram | Records imaging results, focusing on valve conditions and ventricular measurements |
| Other procedures | Other cardiac and noncardiac procedures relevant to the operation are recorded here |
| Perfusion and myocardial protection | Describes intraoperative techniques to protect the heart |
| Postoperative complications | Tracks complications such as reoperations and system failures |
| Discharge details | Summarizes outcomes, including discharge status or causes of death |
| Patient monitoring | Documents postdischarge events to evaluate long-term outcomes and mortality |

The clinical sections with all their variables are important for patient follow-up. Follow-up is standardized at discharge, 30 days after surgery, and 1 year after surgery. Additional follow-up points are added if adverse events occur.

Responsibility for data entry is defined a priori. Data will be entered primarily by surgical nurses and residents during the perioperative care period. Surgeons validate and sign off on each procedure. A registry coordinator (specialist nurse) supervises to ensure completeness. The role of records coordinator is responsible for monitoring the quality of the data entered and the completeness of the record.

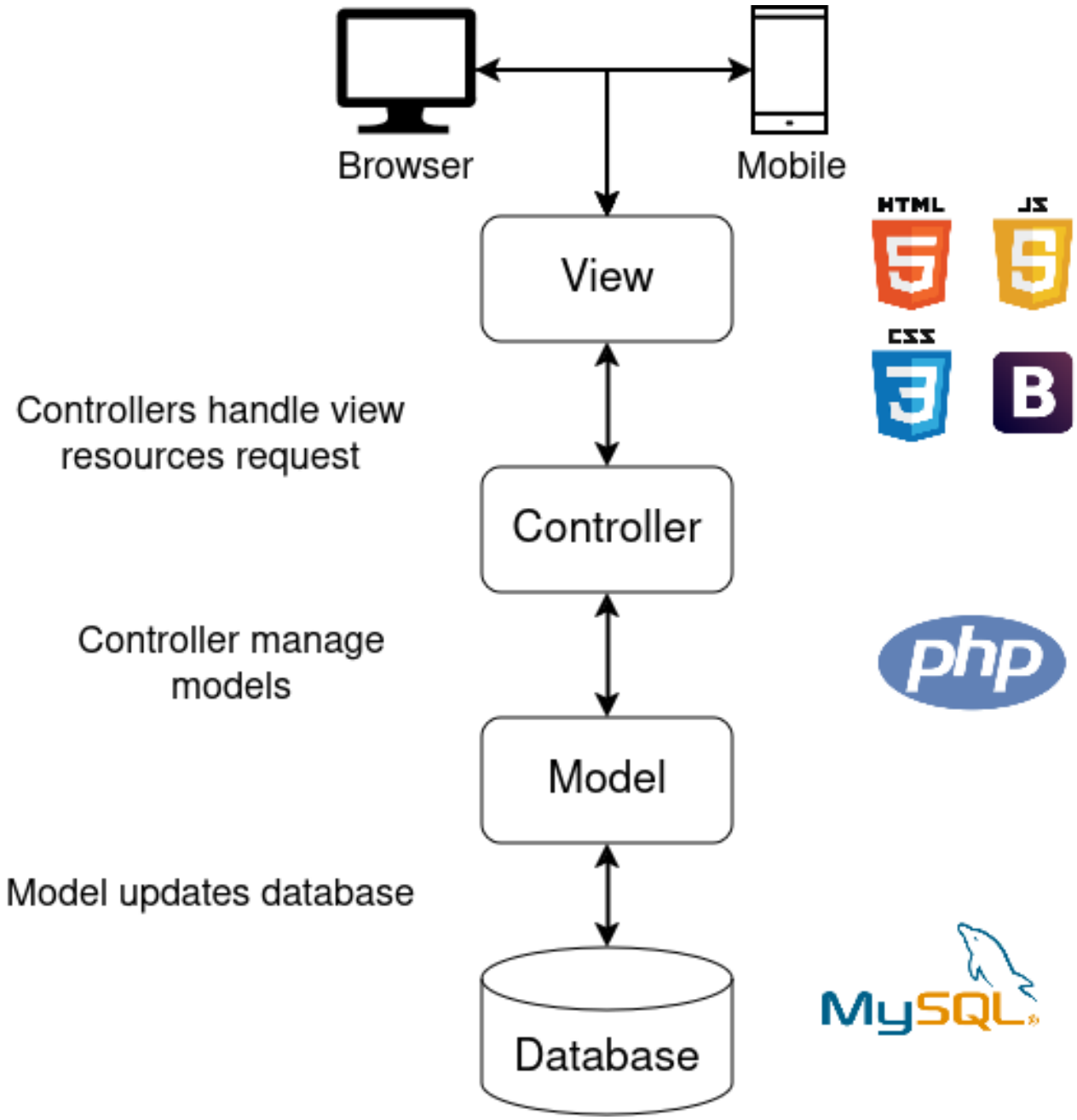
By integrating data from all these categories, the study provides valuable insights into factors influencing surgical outcomes, thereby enhancing our understanding of patient risk profiles and the efficacy of interventions (Multimedia Appendix 1). Minimum dataset for the EACTS-based cardiac surgery registry.

## Software Architecture Design

A web platform based on model-view-controller architecture was developed to ensure modularity, scalability, and ease of maintenance [16]. MySQL 5.6 [17] was used for data persistence, and PHP 5.4 [18] was used to implement the controller and model layers, facilitating efficient server-side scripting and database interaction [19]. The display layer was built using HTML, CSS, JavaScript, and Bootstrap 3, ensuring a dynamic and user-friendly interface with an adaptive design for accessibility across various devices (Figure 2) [20].

The development process began with the design of the database model, followed by the implementation of the server-side components, including security protocols. Responsiveness and usability were prioritized in the client interface. After initial development, a 6-month beta testing phase was conducted at the Guillermo Grant Benavente Hospital in Concepción, a leading cardiology center in Chile. The results of this phase served as a guide to refine the platform and prepare it for production deployment.

**Figure 2.** Model-View-Controller architecture for the cardiac surgery platform.




### Hospital Inclusion Criteria

The platform is currently being used at Guillermo Grant Benavente Hospital in Concepción, Chile. To increase coverage and include more centers, the following inclusion criteria must be satisfied.

- Perform high-complexity cardiac surgeries.
- Have dedicated surgical teams with stable caseloads.
- Commit to institutional agreements ensuring data quality, anonymization, and adherence to ethical and legal requirements.

### Ethical Considerations

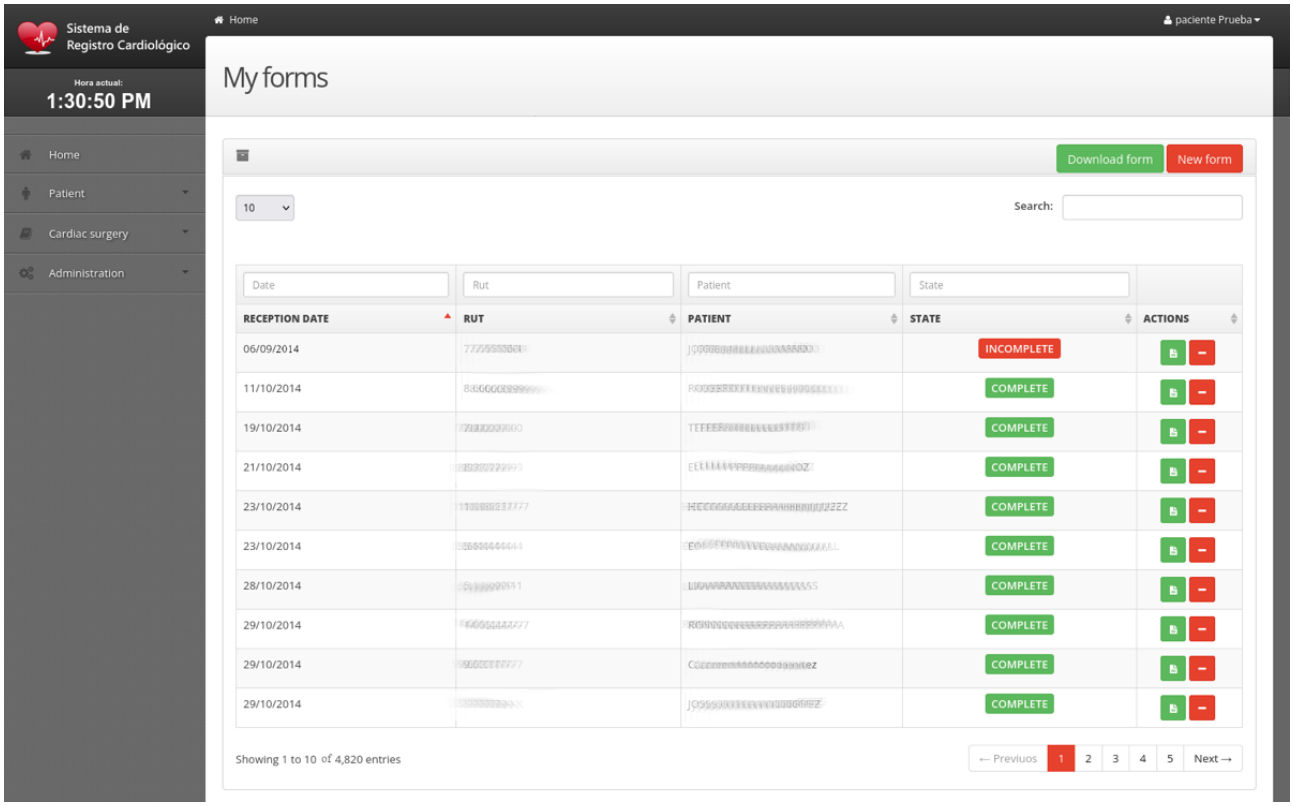
The project and use of the registry were reviewed and approved by the Scientific Ethics Committee of the Faculty of Medicine at Universidad de Concepción (CEC 16/2024). For the purpose of this paper, all data were anonymized.

## Results

### Overview

We created a web platform for systematically recording clinical data from adult cardiac surgeries at the main hospital in south-central Chile (Figure 3). The platform enables efficient data storage, management, querying, and visualization. Since its launch in 2014, it has become the most important registry of complex cardiac procedures in Chile and Latin America [14], documenting more than 4800 surgeries in the past decade (Figure 3), considering 54% (2592/4800) CABG, 34% (1632/4800) valve surgeries, and 12% (576/4800) other complex procedures.

**Figure 3.** Screenshot of the cardiac surgery registry section. The platform facilitates the registration and management of cardiac surgeries performed on patients from various locations in south-central Chile.




The platform is accessible from anywhere with an internet connection and is compatible with various devices, including desktop computers and mobile devices. Its bilingual interface, available in both Spanish and English, ensures accessibility to a wide range of users. Aggregated datasets for advanced analysis, supporting clinical operations. A module dedicated to “Statistics and Graphs” provides predefined reports for immediate use (Figure 1). At the same time, an integrated export tool allows users to extract data.

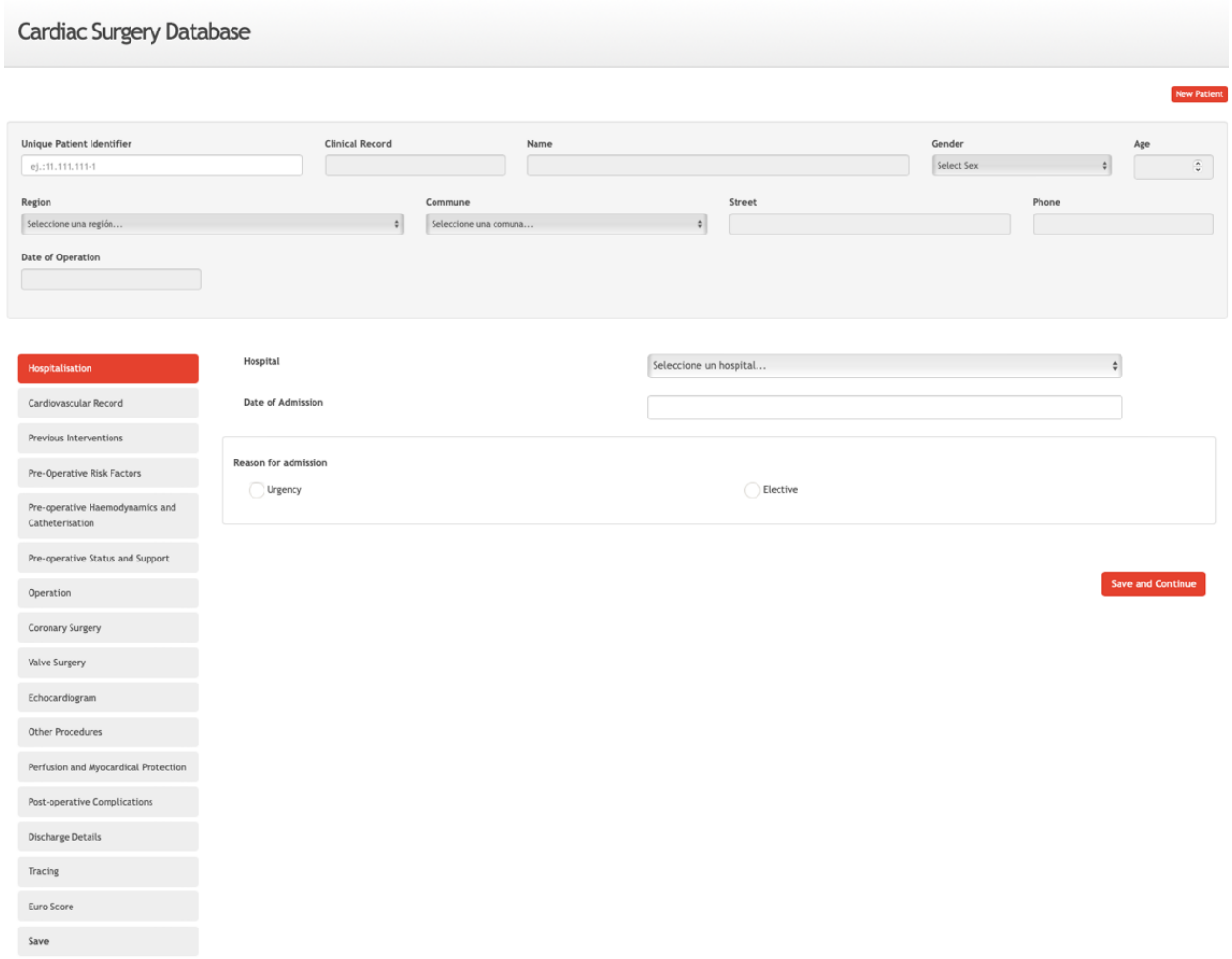
The platform is designed to capture more than 160 structured data points spanning preoperative, intraoperative, and postoperative stages, categorized into 15 sections for comprehensive data collection and streamlined analysis (Figure 4). These sections include patient demographics, preoperative conditions, risk factors, procedural details, postoperative outcomes, discharge information, and follow-up data (Table 2 and Multimedia Appendix 1).

The preoperative stage begins with the hospitalization section, which records patient identifiers and procedural details. The cardiovascular history section captures previous conditions, such as angina and heart failure, while the previous interventions section documents earlier angioplasties and surgeries. Sections on risk factors, hemodynamics, and catheterization status support and detail comorbidities, clinical status, and supportive measures.

The intraoperative section records details of the intervention, including the participating professionals, the reasons for surgery, and the types of procedures performed. Specialized sections for coronary, valve, and echocardiograms provide specific insights, while additional sections cover cardiac and noncardiac procedures, as well as myocardial protection.

The postoperative complications section captures postoperative data, discharge details, patient monitoring, and follow-ups. This structured design ensures standardized and detailed data collection, supporting clinical operations and research initiatives.

**Figure 4.** Screenshot of information recording sections from any internet-connected device.

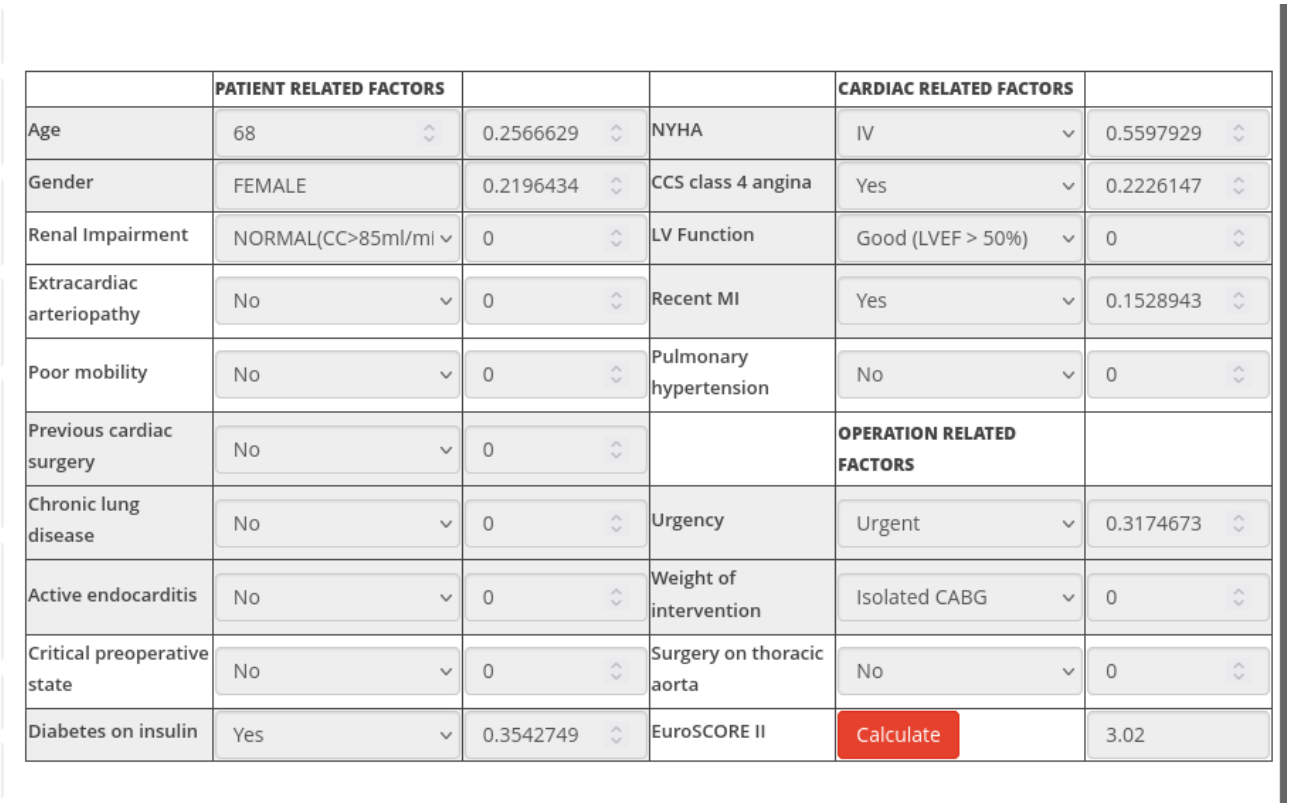



## The EuroSCORE II Module

The EuroSCORE II is a risk prediction model used to estimate the probability of mortality in patients undergoing cardiac surgery [21]. It was developed as an updated version of the original EuroSCORE to reflect contemporary clinical practices and improve accuracy. Widely used in clinical and research settings, EuroSCORE II helps health care professionals assess surgical risks, guide decision-making, and benchmark institutional performance [9,21].

The indicator is derived from factors grouped into 3 main categories: patient-related factors, cardiac-related factors, and operation-related factors (Figure 5). Each factor is assigned a predefined score based on acceptable values, which contributes to the overall risk assessment.

**Figure 5.** Screenshot of the European System for Cardiac Operative Risk Evaluation II module that emphasizes how each factor contributes to the total risk score. CABG: coronary artery bypass grafting; CCS: Canadian Cardiovascular Society; EuroSCORE II: European System for Cardiac Operative Risk Evaluation; LV: left ventricle; LVEF: left ventricular ejection fraction, MI: myocardial infarction; NYHA: New York Heart Association.




To determine the final EuroSCORE II value, all individual scores are summed and applied to a logistic function, which adjusts the raw score into a risk percentage. This process provides an accurate and standardized surgical risk assessment, aiding clinical decision-making and patient counseling.

Each factor is assigned a predefined weight or coefficient based on its relative impact on mortality risk, as determined through large-scale statistical analysis. For example, age contributes progressively higher weights as it increases beyond 60 years, and pre-existing conditions such as renal dysfunction or severe comorbidities significantly increase the risk score.

The cumulative score is then calculated by summing the weighted contributions of all factors. This total score is applied to a logistic regression formula to translate it into a percentage probability of mortality:

$$ESII\% = \frac{e^{\beta 0+\beta 1x1+\beta 2x2+\ldots+\beta nxn}}{1+e^{\beta 0+\beta 1x1+\beta 2x2+\ldots+\beta nxn}} \times 100\%$$

Where:

- e is the base of the natural logarithm
- β0 is the intercept term.
- β1,β2,β3….βn are the regression coefficients for each factor.
- x1,x2,x3….xn are the weighted values of the patient's factors (Figure 5).

## Discussion

### Principal Findings

We present a proposal for the systematic registration and analysis of cardiac surgeries in Chile, aiming to collect clinical data and stratify risk. To address this, we developed a web platform tailored to local needs, offering an intuitive and user-friendly tool. In addition, the platform's alignment with international standards, such as the EACTS guidelines, and the integration of the EuroSCORE II risk model underscore its potential to support both clinical and research applications.

Currently, the registry is implemented and actively used at Guillermo Grant Benavente Hospital, one of the largest reference centers in Chile. Nevertheless, formal agreements are underway with 2 additional cardiac surgery centers (Hospital San Juan de Dios de Curicó and Hospital Clínico Regional de Antofagasta) to expand its use starting in 2025. The long-term vision, discussed with the Ministry of Health, is to scale it as a national platform.

### Platform Implementation and Usability

The platform successfully implemented more than 160 data points, structured across 15 clinical sections, distributed among preoperative, intraoperative, and postoperative variables. Its modular and scalable architecture facilitated integration into clinical workflows while ensuring accessibility through a bilingual interface compatible with various devices. These features have been instrumental in promoting its adoption and ease of use by health care professionals in multiple settings.

### Future Work for the Platform and Risk Assessment

Future development should incorporate artificial intelligence algorithms and advanced statistical methodologies to enhance the platform's impact and improve the accuracy of current risk models [22]. These approaches could improve the predictive accuracy of risk assessments by accommodating complex, nonlinear relationships between variables. In addition, expanding the platform's interoperability with other health information systems would facilitate broader data sharing and benchmarking, thereby aligning local practices with international standards. Additional efforts should be made to engage users in continuous feedback loops to refine the platform's functionality and usability.

The implementation of this platform helps address the challenges associated with fragmented data and limited risk stratification capabilities in Chile. With 4800 records, it is currently the most significant database (considering only 1 center) in Chile and Latin America [1,14].

The collection of data and calculation of the EuroSCORE II will allow the validation of this risk scale in a Latin American population. The differences that can be observed would eventually allow this scale to be adjusted or calibrated to this population. Using a risk scale adjusted to the population to which it is applied will enable better clinical decision-making for our patients.

Based on recent discussions with the Chilean Ministry of Health, it can be extended to the rest of the country, enabling standardized and centralized data collection and laying the foundation for evidence-based improvements in surgical quality and patient safety [2,11]. In addition, risk models such as EuroSCORE II should be recalibrated for the Chilean population to provide accurate and actionable information for informed clinical decision-making.

### Conclusions

This study presents the first Chilean web-based platform for collecting cardiac surgery data, addressing the need for systematic documentation of highly complex procedures. The platform has registered more than 4800 surgeries, encompassing 160 clinical variables. This registry aims to support detailed data analysis and improve surgical planning, resource allocation, and risk assessment by integrating the EuroSCORE II module.

The proposed registry platform is a substantial contribution to clinical centers, and future efforts should focus on improving interoperability and integrating advanced analytics to enable scalability on a national scale.

This initiative demonstrates the potential of biomedical informatics, particularly electronic registry systems, to improve health outcomes, align with international standards, and inform evidence-based public health policies in Chile.

### Acknowledgments

This work was supported by grants from Centro Nacional de Sistemas de Información en Salud (CTI230006) and STARTUP CIENCIA Agencia Nacional de Investigación y Desarrollo de Chile (SUC230035)

### Conflicts of Interest

None declared.

Multimedia Appendix 1
Minimal dataset for record cardiac surgeries based on European Association for Cardio-Thoracic Surgery.
[DOCX File, 38 KB - cardio_v9i1e70147_app1.docx ]

### References

1. Schiller W, Gummert JF. Databases in cardiac surgery. In: Cardiac Surgery: Operations on the Heart and Great Vessels in Adults and Children: Springer; 2017:95-108. [doi: 10.1007/978-3-662-52672-9_5]
2. Hickey GL, Grant SW, Cosgriff R, et al. Clinical registries: governance, management, analysis and applications. Eur J Cardiothorac Surg 2013 Oct;44(4):605-614. [doi: 10.1093/ejcts/ezt018] [Medline: 23371972]
3. Wyler von Ballmoos MC, Kaneko T, Iribarne A, et al. The society of thoracic surgeons adult cardiac surgery database: 2023 update on procedure data and research. Ann Thorac Surg 2024 Feb;117(2):260-270. [doi: 10.1016/j.athoracsur.2023.11.016] [Medline: 38040323]
4. López Menéndez J, Polo López L, Silva Guisasola J, Centella Hernández T. Cirugía cardiovascular en España en el año 2016. Registro de intervenciones de la Sociedad Española de Cirugía Torácica-Cardiovascular [Article in Spanish]. Cirugía Cardiovascular 2017 Nov;24(6):381-397. [doi: 10.1016/j.circv.2017.09.005]
5. Beckmann A, Meyer R, Lewandowski J, Markewitz A, Gummert J. German heart surgery report 2020: the annual updated registry of the German society for thoracic and cardiovascular surgery. Thorac Cardiovasc Surg 2021 Jun;69(4):294-307. [doi: 10.1055/s-0041-1730374] [Medline: 34176107]
6. Dimagli A, Guida G, Sinha S, et al. Surgical outcomes of post‐infarct ventricular septal defect repair: insights from the UK national adult cardiac surgery audit database. J Card Surg 2022 Apr;37(4):843-852. [doi: 10.1111/jocs.16178] [Medline: 34910324]

7. Tran L, Williams-Spence J, Shardey GC, Smith JA, Reid CM. The Australian and New Zealand society of cardiac and thoracic surgeons database program - two decades of quality assurance data. Heart Lung Circ 2019 Oct;28(10):1459-1462. [doi: 10.1016/j.hlc.2019.03.002] [Medline: 30962063]
8. Committee for Scientific Affairs, The Japanese Association for Thoracic Surgery, Shimizu H, Okada M, et al. Thoracic and cardiovascular surgeries in Japan during 2018: annual report by the Japanese Association for Thoracic Surgery. Gen Thorac Cardiovasc Surg 2021 Jan;69(1):179-212. [doi: 10.1007/s11748-020-01460-w] [Medline: 33090365]
9. Head SJ, Howell NJ, Osnabrugge RLJ, et al. The European Association for Cardio-Thoracic Surgery (EACTS) database: an introduction. Eur J Cardiothorac Surg 2013 Sep;44(3):e175-e180. [doi: 10.1093/ejcts/ezt303] [Medline: 23786918]
10. EACTS adult cardiac database file specification. : The European Association for Cardio-Thoracic Surgery; 2018 URL: https://www.eacts.org/wp-content/uploads/2018/03/EACTS-ACD-File-Specification-01-03-18.pdf [accessed 2025-10-21]
11. Dawson LP, Biswas S, Lefkovits J, et al. Characteristics and quality of national cardiac registries: a systematic review. Circ Cardiovasc Qual Outcomes 2021 Sep;14(9):e007963. [doi: 10.1161/CIRCOUTCOMES.121.007963] [Medline: 34517724]
12. Sommerville I. Software Engineering, 9th edition: Pearson; 2011. URL: https://epage.pub/doc/software-engineering-9ed-354prljv4e [accessed 2025-10-22]
13. EACTS is the home for the global cardiothoracic surgical community. European Association for Cardio-Thoracic Surgery (EACTS). URL: https://www.eacts.org/ [accessed 2024-11-25]
14. Cerda-Núnez C, Yánez-Lillo J, Seguel E, Guinez-Molinos S. Performance of EuroSCORE II in Latin America: a systematic review. Rev Med Chil 2022 Apr;150(4):424-430. [doi: 10.4067/S0034-98872022000400424] [Medline: 36155751]
15. Satapathy A, Livingston J. A comprehensive survey on SSL/ TLS and their vulnerabilities. IJCA 2016;153(5):31-38. [doi: 10.5120/ijca2016912063]
16. Taibi D, Lenarduzzi V, Pahl C. Architectural patterns for microservices: a systematic mapping study. Presented at: 8th International Conference on Cloud Computing and Services Science; Mar 19-21, 2018; Funchal, Madeira, Portugal p. 221-232. [doi: 10.5220/0006798302210232]
17. MySQL 5.6 reference manual including mySQL NDB cluster 7.3-7.4 reference guide. : Oracle; 2023 Feb URL: https://downloads.mysql.com/docs/refman-5.6-en.pdf [accessed 2025-10-21]
18. PHP 5.4.0 release announcement. PHP. URL: https://www.php.net/releases/5_4_0.php [accessed 2024-11-25]
19. Williams HE, David L. Web Database Applications with PHP and mySQL: O'Reilly; 2004. URL: https://www.oreilly.com/library/view/web-database-applications/0596005431/ [accessed 2025-10-21]
20. Bootstrap is the most popular HTML, CSS, and JS framework for developing responsive, mobile first projects on the web. Bootstrap. URL: https://getbootstrap.com/docs/3.3/ [accessed 2024-11-25]
21. Nashef SAM, Roques F, Sharples LD, et al. EuroSCORE II. Eur J Cardiothorac Surg 2012 Apr;41(4):734-744. [doi: 10.1093/ejcts/ezs043] [Medline: 22378855]
22. Miles TJ, Ghanta RK. Machine learning in cardiac surgery: a narrative review. J Thorac Dis 2024 Apr 30;16(4):2644-2653. [doi: 10.21037/jtd-23-1659] [Medline: 38738250]


## Abbreviations

**CABG:** coronary artery bypass grafting
**EACTS:** European Association for Cardio-Thoracic Surgery
**EuroSCORE:** European System for Cardiac Operative Risk Evaluation


*Edited by J Rivers; submitted 16.12.24; peer-reviewed by A Rouhi, F Harig, Y Jiang, Z Zheng; revised version received 12.09.25; accepted 10.10.25; published 11.11.25.*

*Please cite as:*
*Guinez-Molinos S, Seguel E, Gonzalez J, Castillo B*
*Web-Based Platform for the Chilean Cardiac Surgery Registry: Algorithm Development and Validation Study*
*JMIR Cardio 2025;9:e70147*
*URL: https://cardio.jmir.org/2025/1/e70147*
*doi:10.2196/70147*

# The Impact of Digital Intervention Messages Targeting Users With High Blood Pressure Events: Retrospective Real-World Study

Yifat Fundoiano-Hershcovitz[1], PhD; Inbar Breuer Asher[1], MPH; Marilyn D Ritholz[2], PhD; David L Horwitz[3], MD; Omar Manejwala[1], MD; Claudia Levi[1], NP; Pavel Goldstein[4], PhD

[1]DarioHealth, Hatarshish 5, North Industrial Park, Caesarea, Israel

[2]Joslin Diabetes Center, Harvard Medical School, Boston, MA, United States

[3]DLH Biomedical Consulting, Las Vegas, NV, United States

[4]School of Public Health, University of Haifa, Haifa, Israel

**Corresponding Author:**
Yifat Fundoiano-Hershcovitz, PhD
DarioHealth, Hatarshish 5, North Industrial Park, Caesarea, Israel

## Abstract

**Background:** Effective hypertension management, particularly through self-care strategies, remains a significant public health challenge. Despite widespread awareness, only approximately 1 in 5 adults achieves adequate blood pressure (BP) control. There is a growing need for scalable digital health interventions that enhance awareness, support behavioral change, and improve clinical outcomes. However, real-world evidence evaluating the impact of such interventions on BP levels and their underlying mechanisms is limited.

**Objective:** This study aimed to evaluate the effectiveness of a digital intervention using data-driven nudges on monthly average BP levels. Specifically, we assessed changes in BP before and after the intervention and examined whether these changes differed compared to a control group in a high BP cohort and a normal BP cohort.

**Methods:** In this retrospective, real-world cohort study, we analyzed two user cohorts from a digital health platform: (1) individuals with high BP readings and (2) individuals with normal BP readings. Participants who received a digital intervention were propensity score–matched to users who did not receive the intervention, based on demographic and clinical variables. Monthly average BP and the proportion of high readings were assessed 3 months before and after the intervention. A piecewise mixed-effects model was used to evaluate BP trajectories, and simple slope analysis assessed the interaction between the outcomes and the groups, as well as the moderating effect of lifestyle activities on systolic blood pressure (SBP).

**Results:** In total, 408 users were included in the study. In the high BP cohort (n=296), the intervention group showed a significant decrease in the monthly average SBP after the intervention (B=–2.09; *P*<.001), while the control group showed a smaller reduction (B=–1.06; *P*=.007). Additionally, users reporting higher lifestyle activity levels experienced a greater reduction in SBP (B=–5.27; *P*<.001). In the normal BP cohort (n=112), the intervention group maintained stable BP levels after the intervention (B=–0.39; *P*=.27), while the control group exhibited a significant increase in BP levels (B=0.69; *P*=.03).

**Conclusions:** Data-driven nudges delivered via a digital health platform were associated with improved BP outcomes among individuals with high BP levels and helped maintain BP stability among those with normal BP levels. These findings reinforce the integration of personalized digital interventions into hypertension management and highlight the potential role of positive messaging, behavioral engagement, and user empowerment in improving long-term outcomes.

***(JMIR Cardio 2025;9:e76275)*** doi:10.2196/76275

**KEYWORDS**

digital health intervention; blood pressure; hypertension; retrospective study; lifestyle behaviors

## Introduction

High blood pressure (BP) is a major public health challenge [1], but it has also been identified as the leading preventable risk factor for premature death [2]. Hypertension has been diagnosed in approximately 1.4 billion (31%) adults aged 30 to 79 years worldwide [3] and 108 million adults in the United States [4]. The prevalence of hypertension is rising globally owing to the aging of the population and an increase in the exposure to lifestyle risk factors [3] such as an unhealthy diet and sedentary lifestyle [5]. Additionally, the implementation of new guidelines in 2017, which lowered the diagnostic threshold for hypertension to systolic blood pressure (SBP) ≥130 mm Hg and diastolic blood pressure (DBP) ≥90 mm Hg [6], has led to more people being classified as hypertensive.

In addition, high BP is controlled in only about 1 in 5 adults (21%) [7,8], defined as maintaining SBP below 140 mm Hg and DBP below 90 mm Hg, according to the Joint National Committee (JNC7) guidelines. Furthermore, there are substantial disparities in disease awareness, treatment, and control across different racial and ethnic groups in the United States [9]. Many factors contribute to these inequalities, such as health literacy, socioeconomic status, reduced access to healthy foods, and health literacy [9]. These disparities impact the overall burden of hypertension in the United States and highlight the need for health care strategies to address hypertension across all populations.

Hypertension is a chronic condition that fluctuates over time, with periods of stability interrupted by episodes of elevated pressure. This variability is one of the reasons why hypertension is so challenging to manage. Users may experience fluctuations that complicate diagnosis and long-term management [10]. Monitoring BP regularly is key to minimizing these ups and downs and helping prevent the complications associated with chronic hypertension [11,12].

Treatment and management of hypertension are critically important for the reduction of cardiovascular complications and for the prevention of consequent diseases [13,14]. Despite the proven efficacy of pharmacological treatments [15] and the effectiveness of the variety of nonpharmacological interventions in lowering BP [16], poor BP control remains a pervasive problem. Suboptimal adherence, characterized by the failure to initiate treatment and to persist in therapy in the long term, is a well-recognized factor contributing to the inadequate control of BP in hypertension [17]. Studies have shown that low adherence is associated with reduced therapeutic success, reduced quality of life, and higher treatment costs [18].

Numerous factors can influence adherence in users with hypertension, including older age, lower education levels, potential medication side effects, and insufficient guidance from health care professionals [19]. Additionally, low adherence may stem from the common misconception among users that medication is unnecessary, as hypertension often presents without symptoms [19]. This highlights the importance of user education in managing hypertension, as it empowers them to make informed choices and effectively control risk factors, which can ultimately improve long-term health outcomes.

Health technology has created new opportunities to improve the management and treatment of chronic conditions like hypertension [20], particularly after the COVID-19 pandemic. Self-management education and support have been widely used as strategies aiming to provide users with the appropriate health literacy and skills for the effective, long-term control of hypertension [21]. In recent years, due to the fact that 86% of the global population has access to a smartphone, digital interventions allow a more convenient and accessible form of health care delivery, resulting in effective hypertension self-management [21].

Several studies have assessed the benefits of mobile health (mHealth) in promoting BP self-management and its effectiveness in managing other cardiometabolic conditions [17,22,23]. Digital tools provide a promising, cost-effective, and scalable solution to improve and sustain hypertension outcomes on a large scale. Meta-analyses have demonstrated that mHealth interventions not only reduce BP but also increase the reach, uptake, and feasibility of hypertension management [21]. Digital solutions for hypertension must be evidence-based and effective to reduce its impact on global noncommunicable diseases [24].

Progress in mHealth technology has enabled the design of just-in-time adaptive interventions (JITAIs) [25]. JITAIs have emerged as impactful approaches, providing support for behavior change and hypertension monitoring [26]. This type of intervention enables the delivery of real-time support and provides personalized contextual feedback. Despite the significant advancements in technology and the appeal of JITAIs, many programs have been developed with little empirical evidence, and research has been limited by a lack of evidence regarding effectiveness and sustained engagement [25,26].

Our study performed a retrospective analysis of a digital health platform for hypertension management by integrating a home-use BP-monitoring system with comprehensive data captured through a supportive mobile app for individuals with normal and poorly controlled BP levels. The study aimed to assess how a BP digital intervention with data-driven nudges would influence monthly average BP levels, with measurements of BP 3 months before and after the nudges. We hypothesized that prior to the delivery of the nudges, the 2 groups would exhibit similar BP levels. However, following the delivery of the nudges, the groups would show distinct trajectories in their BP levels. Specifically, we expected that the group receiving the nudges would demonstrate a greater reduction in their BP compared to the group not receiving the nudges. By analyzing the 2-stage period, we aimed to gain a comprehensive understanding of the impact of educative messages and positive feedback, assessing whether changes occurred between the pre- and postintervention phases and whether these changes differed between the 2 groups.

## Methods

### Platform

This study used the Dario Health digital health platform to support the self-management of hypertension within the context of chronic cardiometabolic conditions. The Dario BP monitoring system combines a connected BP monitor with a mobile app (compatible with both Android and iOS devices). The BP-monitoring system measures the SBP, DBP, and pulse rate by using a noninvasive technique in which an inflatable cuff is wrapped around the upper arm. Viewing both pulse rate and BP provides a more comprehensive picture of cardiovascular health. These 2 metrics can impact each other in different ways and offer valuable insights into a user's overall cardiovascular health. For example, an unexpected combination of high or low readings can be a key indicator of underlying health issues. Similarly, a normal BP paired with a low resting pulse rate often indicates strong cardiovascular fitness [27,28]. The BP-monitoring system uses Bluetooth. The BP cuff is paired with the mobile app, and the data are transmitted to the smart mobile device via Bluetooth,

ensuring real-time feedback and 100% data capture. The BP reading is displayed on the mobile app screen. The immediate display of measurements on the smartphone interface supports timely decision-making and enhances engagement.

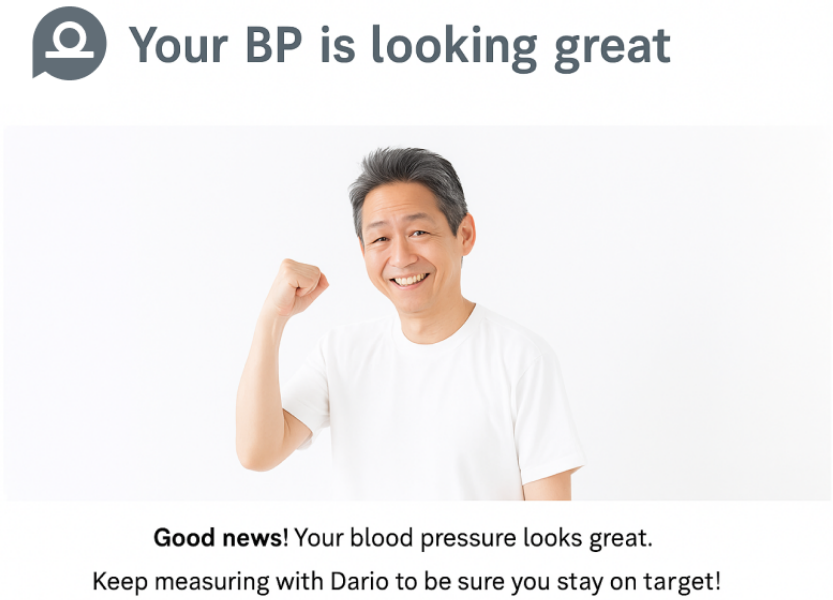
Users can input additional data at the time of measurement, such as the arm used, recent activities (eg, smoking, caffeine intake, and physical activity), and symptoms (eg, stress, dizziness, and headache). This information is securely stored in a digital logbook, with automatic cloud backup enabling further analysis and clinical interventions tailored to individual needs.

The platform's data-driven approach uses real-time nudges designed to improve health outcomes. These messages provide personalized feedback and educational inputs based on BP thresholds aligned with American Heart Association (AHA) guidelines, which provide standardized targets for BP management across all nonpregnant adults, regardless of age or gender [29]. Personalization is achieved through dynamic tailoring of messages based on each user's clinical profile and real-time data trends. The system analyzes users' longitudinal BP patterns and identifies trends, delivering in-app notifications, push messages, emails, and SMS text message alerts with actionable insights. When BP readings indicate specific trends, users receive targeted educational content, motivational messages, and behavioral prompts to encourage adherence to health goals and increase platform engagement. Specifically, the system analyzes users' BP patterns to identify meaningful changes or persistently elevated readings. Messages are then customized according to these trends, offering feedback relevant to the user's current status, such as positive reinforcement for sustained control or motivational nudges when elevated values are detected. For example, a user with persistently high SBP readings may receive targeted lifestyle suggestions (eg, sodium reduction and physical activity prompts), while a user demonstrating improvement may receive encouraging messages reinforcing adherence.

Three independent readings over 3 individual days (nonconsecutive) within a 7-day period will trigger a digital intervention, except for a hypertensive crisis, which causes a trigger for each event. The BP levels embedded in the system are similar to those defined by the AHA [29]: normal BP, SBP <120 mm Hg and DBP <80 mm Hg; elevated BP, SBP 120 - 129 mm Hg and DBP <80 mm Hg; hypertension stage 1, SBP 130 - 139 mm Hg or DBP 80 - 89 mm Hg; hypertension stage 2: SBP ≥140 mm Hg or DBP ≥90 mm Hg; and hypertensive crisis, SBP >180 mm Hg and/or DBP >120 mm Hg.

For normal BP clusters, motivational feedback was delivered via SMS text messages and in-app messages encouraging the user to keep measuring their BP to stay on target. The nudges triggered by normal BP levels included the following: "Well done! Your BP is looking good," "Good news. Your blood pressure looks good," "Keep measuring with Dario to be sure you stay on target!," and "Your BP is looking great." A representative message is presented in Figure 1.

**Figure 1.** Representative intervention message delivered via the Dario app following normal blood pressure readings.



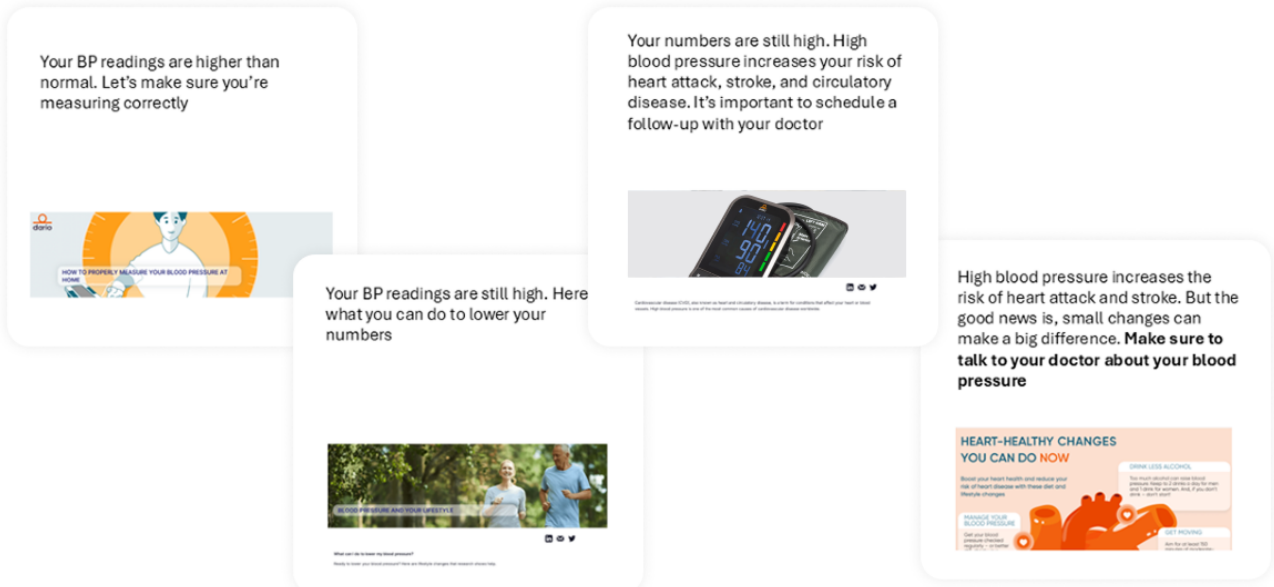

For stage 1 hypertension, users received informational push notifications, SMS text messages, and educational articles through cluster events. Messaging was nonjudgmental in tone and included the following: "Your blood pressure readings are higher than normal, according to your numbers in the Dario App. Are you measuring correctly? Review the steps for measuring at home," "Have you heard about the DASH diet? Link to DASH ref Learn more about the diet plan that has been found to help lower blood pressure," and "Recent measurements show that your blood pressure is still higher than normal. If you're taking your blood pressure medication as directed, it's time to schedule a medication review. Want to make sure you are taking your medications correctly? Review a list of things you need to know."

Educational topics were based on current hypertension management guidelines that recommend, as an integral part of ongoing treatment, the adoption of lifestyle modifications, including a healthy diet, independently of the underlying antihypertensive drug treatment [29]. Research has provided lifestyle recommendations, including a low-sodium, increased-potassium, low-fat diet; maintenance of appropriate body weight; alcohol use reduction; smoking cessation; increased physical activity; and stress reduction [30]. The average impact of each lifestyle change is a decrease of 4 - 5 mm Hg in SBP and a decrease of 2 - 4 mm Hg in DBP; however, a diet low in sodium, saturated fat, and total fat and an increase in the consumption of fruits, vegetables, and grains may decrease SBP by approximately 11 mm Hg [29]. Educational content includes general information on the meaning of BP values, the adoption of a healthy lifestyle or diet, and adherence to medication. Lifestyle changes include moving toward a healthy body weight, performing physical activity of 30 minutes a day, eating a heart-healthy diet, reducing alcohol consumption, quitting smoking, managing stress, and regularly monitoring BP [31].

DASH (Dietary Approaches to Stop Hypertension) studies have shown that diets rich in fruits and vegetables and low in saturated and total fats can both lower the risk of high BP and assist with BP control in people with hypertension. Vegetables and fruits account for approximately half of the BP-lowering effect of the diet. Foods in the DASH diet are rich in minerals such as potassium, calcium, and magnesium. The diet limits foods that are high in sodium. It also limits added sugar and saturated fat, such as that in fatty meats and full-fat dairy products. The standard DASH diet limits salt intake to 2300 mg a day. This amount agrees with the Dietary Guidelines for Americans [32]. A lower-sodium version of the DASH diet restricts sodium to 1500 mg a day. Restricting sodium intake can enhance the BP-lowering effect. While the DASH diet can reduce SBP by 5 - 6 mm Hg, individuals eating the DASH diet in combination with the lowest sodium intake have been reported to achieve a further BP decrease of 7.1 mm Hg [33]. Compliance with prescribed therapies is a pivotal factor in treatment success. According to the World Health Organization, medication adherence can have a more direct impact on clinical outcomes than the specific treatment itself. Multiple factors contribute to adherence levels, stemming from individual, provider, and health care system elements, which often interact with each other. In the United States, only 51% of users adhere to their medication regimen for high BP [34,35]. Through the messages of the intervention, users also receive guidance on proper medication use to enhance adherence.

Representative messages for high reading clusters are presented in Figure 2.

**Figure 2.** Representative intervention messages delivered via the Dario app following high blood pressure readings.




## Measures

The monthly average BP level (SBP and DBP), which was defined as the mean of a user's BP measurements taken over a 30-day period from the first nudge message, was used as a core outcome metric. Monthly aggregation was specifically chosen based on several evidence-based considerations: (1) it aligns with clinical practice guidelines that recommend assessing BP control trends over weeks to months rather than days [29]; (2) it provides sufficient statistical power to detect clinically meaningful changes while reducing noise from day-to-day physiological variability, which can be as high as 10 - 15 mm Hg due to factors unrelated to intervention effects [12]; (3) it captures the sustained behavioral change trajectory that digital interventions aim to achieve, as behavior modification typically requires 3 - 4 weeks to establish new habits [36]; and (4) it corresponds to the temporal scale at which medication adjustments and lifestyle interventions are evaluated in clinical trials, facilitating comparison with existing literature.

The monthly high-reading percentage, which was defined as the monthly number of high readings (SBP ≥140 mm Hg and DBP ≥90 mm Hg) divided by the monthly number of all BP measurements taken over a 30-day period from the first nudge message, was used as another core outcome metric. The mobile

platform collected the following medical and sociodemographic information (by self-report) for each user: gender, age, weight, BMI, physical activity level, stress level (0 [no stress] to 10 [very stressed]), alcohol consumption (number of drinks per week), smoking (0 [never] to 3 [yes]), and comorbidities (eg, baseline hypertension [based on the first 30-day measurements on the platform], diabetes, high lipid levels, chronic kidney disease, cardiovascular disease, and cancer). Independent variables included digital engagement, such as the number of monthly BP measurements, and lifestyle activities (operationalized as the sum of meal logs, carbohydrate intake, calories burned, and recipe finder activities in each month).

## Study Population

A retrospective data study was performed on the Dario database. Individuals who used the Dario platform between 2019 and 2024 were considered. The users purchased the device via a direct-to-consumer channel. The study analyzed 2 cohorts of users who received a digital intervention (intervention group): one cohort of users who had high BP readings and another cohort of users who had normal BP readings [29]. The inclusion criterion for the intervention group (both cohorts) was measurement of BP using the Dario Health platform during the years 2021 - 2024, for a minimum of 2 months (1 month prior to the first nudge message and 1 month after). To establish a control group, we selected users from the existing Dario database who never received a nudge between August 2019 and May 2020, and who experienced the same BP events. We applied the propensity score matching procedure to ensure comparability. The events that captured and triggered a digital intervention included high BP events (defined as SBP ≥130 mm Hg or DBP ≥80 mm Hg) occurring 4 times on different days within a 7-day period in the high BP cohort, and normal BP events (defined as SBP <120 mm Hg or DBP <80 mm Hg) occurring 4 times on different days within a 7-day period in the normal BP cohort.

## Ethical Considerations

All data used for the analysis were anonymized before extraction for this study. The study received an exemption from Ethical and Independent Review Services (a professional review board), which issued the institutional review board exemption (number: 18,032 - 07#) [37]. The users who participated in the study were provided with a Terms of Use document mentioning the legally valid consent of the end user for the company to collect and access their information. The use of the app, site, or services is deemed to constitute user consent that is legally bound by the Terms of Use and the Privacy Policy. The Terms of Use do not specify an option for users to opt out of the use of their deidentified data for research purposes while continuing to use the service. The current Terms of Use can be accessed on DarioHealth [38]. No compensation was provided to users.

## Study Design

The aim of our study was to evaluate the impact of digital interventions on BP levels and assess their relative contribution to BP levels. We conducted 2 separate parallel analyses: one for the high BP cohort and another for the normal BP cohort. For the digital intervention group, it was crucial to establish a clear starting point for the intervention to assess its effects accurately.

All users in the intervention and control groups had access to the same Dario platform features, including BP monitoring, educational materials, and lifestyle tracking tools. The sole difference was that the intervention group received automated, personalized nudge messages triggered by their BP patterns, while the control group did not receive these triggered messages despite experiencing the same BP events.

This approach ensured that any observed differences in outcomes would be attributable to the digital intervention itself, rather than temporal factors or external influences. Using this approach, we enhanced the internal validity of the study. This allowed us to isolate the effect of the digital intervention from that of other variables and assess the impact on BP levels more accurately.

## Propensity Scores: Causal Inference

Propensity score matching was used in this study to address potential confounding factors and enhance the comparability of the intervention and control groups. The rationale behind using propensity score matching lies in its ability to reduce bias and mimic the randomization process, thereby facilitating causal inference in observational studies [39].

## Missing Data Assessment and Handling

Prior to the analysis, we conducted a comprehensive assessment of missing data patterns and mechanisms. The Little missing completely at random test [40] was performed, and there was no evidence against the missing completely at random assumption ($\chi^2_{11}$=16.438; *P*=.13). Missing data were present only in baseline covariates used for propensity score matching, while outcome variables (BP measurements) had almost complete data at the month-aggregated level.

The missing data in both datasets were imputed using the Multivariate Imputation by Chained Equations (MICE) algorithm with the “mice” package in R (R Project for Statistical Computing). Specifically, the method used for imputation was predictive mean matching, with 5 imputed datasets generated (m=5) and a maximum of 50 iterations (maxit=50) for convergence. This approach allows for the generation of plausible values for missing data based on observed relationships in the dataset, ensuring that the imputed data preserve the underlying statistical structure.

Propensity score matching was used to estimate intervention effects while removing the bias created by the intervention covariates [41] and forming matched sets of treated and untreated individuals who share a similar value of the propensity score. The goal of propensity score matching is to simulate the conditions of a randomized controlled trial, creating a balance between groups in the distributions of covariates.

The matching [41-44] for the high BP cohort was based on the following sociodemographic and clinical parameters: age, gender, weight, baseline hypertension as a binary variable (based on the first 30-day measurements on the platform), number of comorbidities, diabetes type (type 2, prediabetes, or others),

mean number of BP measurements, average SBP and DBP in the months before the messages were sent, and average SBP and DBP as well as mean high-reading percentage in the month the messages were sent. The matching for the normal BP cohort was based on the following sociodemographic and clinical parameters: age, gender, weight, baseline hypertension as a binary variable (based on the first 30-day measurements on the platform), number of comorbidities, diabetes type (type 2, prediabetes, or others), number of BP measurements, average SBP and DBP in the months before the messages were sent, and the first SBP and DBP measurements in each individual.

We used nearest-neighbor matching without replacement to ensure optimal balance between groups while maintaining interpretability. This approach was chosen over other methods (eg, full matching and weighting) for the following reasons: (1) it provides the most intuitive comparison between matched pairs, (2) it performs well when the overlap in propensity scores is substantial, and (3) it allows for straightforward assessment of covariate balance. A caliper width of 0.1 SDs of the logit of the propensity score was selected based on empirical evidence from Austin [43], who demonstrated that calipers of 0.2 SD eliminate approximately 99% of the bias due to measured confounders. We chose a more stringent caliper of 0.1 SD to further minimize potential bias while monitoring the trade-off with sample size reduction. Postmatching balance was assessed using standardized mean difference (SMD) for all covariates, with SMD <0.1 considered indicative of good balance [42]. As shown in Figures 3 and 4, all covariates achieved SMD <0.1 after matching, confirming the effectiveness of our matching procedure. Additionally, we examined the distribution of propensity scores before and after matching to ensure adequate overlap.

**Figure 3.** Plot presenting the efficacy of the matching procedure for balancing the high blood pressure (BP) cohort. Distance refers to the standardized difference in the propensity scores between the treatment and the control after matching. A caliper of 0.1 SD is reached for all parameters.

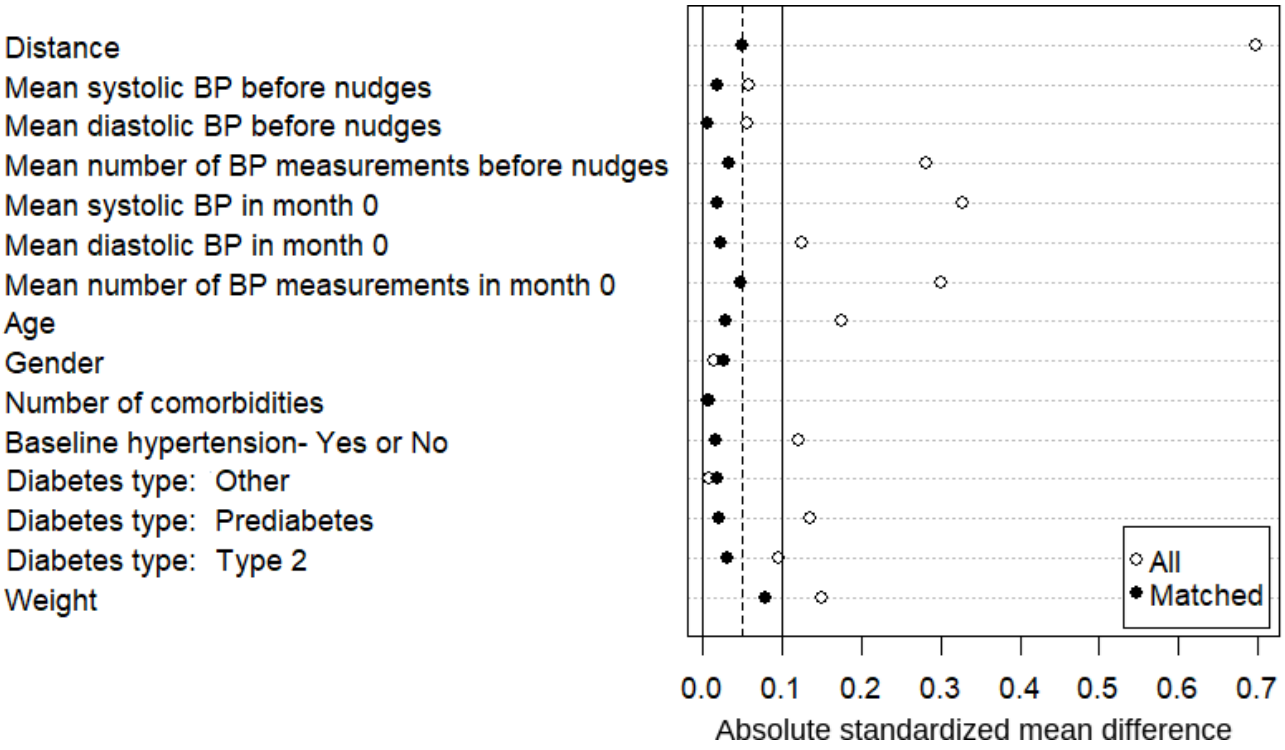
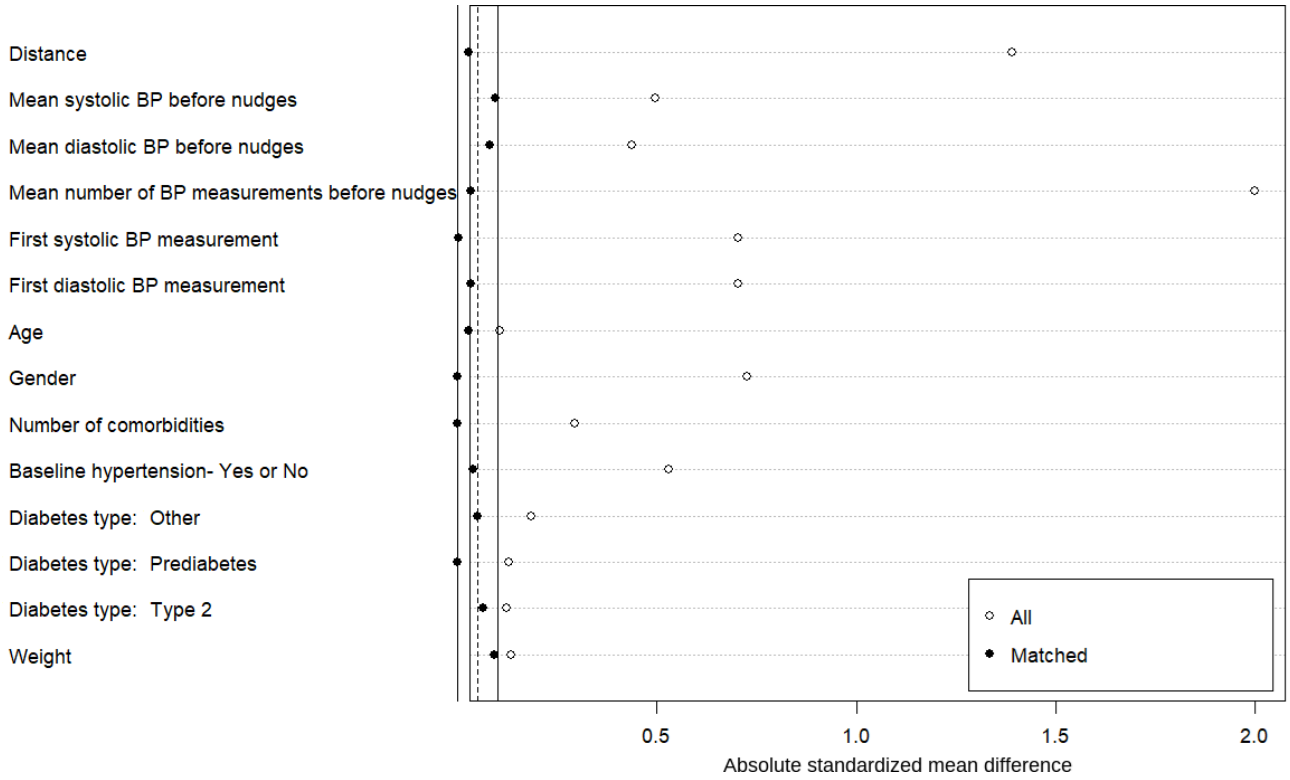


**Figure 4.** Plot presenting the efficacy of the matching procedure for balancing the normal blood pressure (BP) cohort. Distance refers to the standardized difference in the propensity scores between the treatment and the control after matching. A caliper of 0.1 SD is reached for all parameters.




In this study, the propensity scores were calculated for each participant using the "matchit()" function from the R package *matchit*, and the distance metric used was based on logistic regression using a 1:1 ratio between the 2 study groups: 148 users in each group from the high BP cohort and 56 users in each group from the normal BP cohort. We applied nearest-neighbor matching with a caliper width of 0.1 SDs of the propensity score using logistic regression of the intervention on the covariates. Participants without suitable matches were excluded from the analysis. Figures 3 and 4 present the efficacy of the matching procedures for balancing the groups.

## Analytic Approach

A classical linear longitudinal model assumes a single-slope growth pattern for changes in an outcome variable across time. In contrast, piecewise - based mixed - effects models allow flexibility in the modeling of variable change trajectories across time [45]. Here, a piecewise mixed-effects model assessed differences in SBP and DBP in 2 segments: before and after the nudge messages were sent. The piecewise model allowed the data to exhibit different linear trends over their different regions.

In the intervention group, user measurements were centered around receiving the first nudge message in the high BP cohort and the normal BP cohort. In the control group, user measurements were centered around the time of the cluster of events that were the same as in the intervention group. For the high BP cohort, data from 3 months prior to and 3 months following the intervention were included in the analysis. For the normal BP cohort, the analysis included data from 3 months prior to and 6 months following the intervention to assess the sustainability of normal BP events over the longer term.

A piecewise-based mixed-effects model was fitted, modeling temporal changes of the monthly average SBP and DBP for the 2 groups in each cohort (high and normal BP). For the high BP cohort, the monthly average of the high-reading percentage was also modeled. The piecewise cutoff point for the model was set for month 0 (the month that the messages were sent), assuming a change in the time-related monthly average BP between the groups by the included interaction terms between the 2 time trajectories and the groups. Another piecewise mixed-effects model was evaluated only in the high BP cohort to model the effect of the interaction between the 2 time periods and lifestyle activities on the monthly average SBP. We compared three random effect specifications: (1) random intercepts only, (2) random slopes and intercepts, and (3) uncorrelated random effects. Model comparison using the Bayesian information criterion supported the random intercept–only structure that was used in the models. All piecewise mixed-effects models were fitted using restricted maximum likelihood estimation via the lme4 package in R version 4.3.1. Convergence was assessed using gradient tolerance <0.002 and confirmed by positive definite Hessian matrices. Model assumptions were validated through Diagnostics for Hierarchical Regression Models–simulated residual diagnostics and examination of random effect distributions. Next, we used a simple slope analysis to interpret the interaction between the outcomes and the groups. The same statistical framework was applied to test the moderating effect of lifestyle activities on SBP in the high BP cohort. All coefficients from the piecewise mixed-effects models represent the rate of change in the outcome variable per month. For BP outcomes (SBP and DBP), coefficients are expressed in mm Hg per month, indicating the average monthly change during each time period.

## Results

### Users

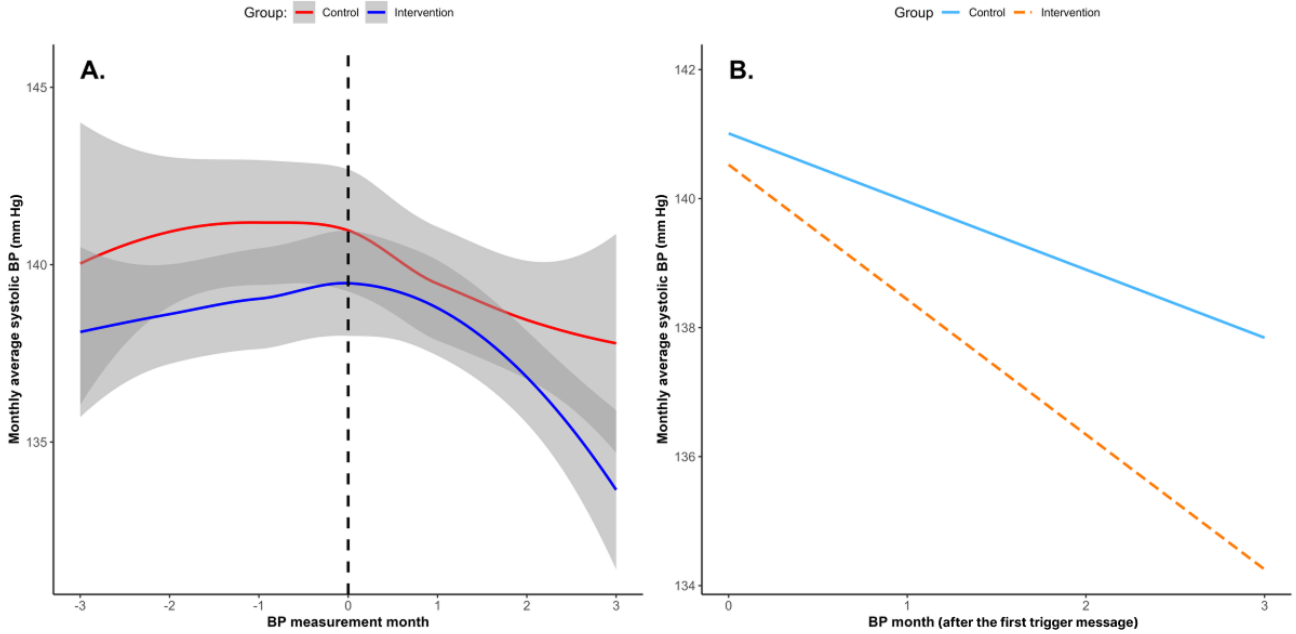
In total, 408 users were included in the study. The high BP cohort had 296 users, including 186 (62.8%) men and 110 (37.2%) women. In this cohort, the average user age was 63.9 (SD 10.0) years, and the average BMI was 32.2 (SD 6.9). Of the 296 users, 196 (66.2%) reported having one or more comorbidities. The normal BP cohort had 112 users, including 30 (26.8%) men, 49 (43.8%) women, and 33 (29.4%) others. In this cohort, the average user age was 55.1 (SD 11.6) years, and the average BMI was 32.0 (SD 10.1). Of the 112 users, 37 (33.0%) reported having one or more comorbidities.

Sensitivity analysis comparing complete case analysis (n=237) with imputed results demonstrated robust findings. In complete cases, the intervention group showed a decrease in monthly SBP of −1.73 mm Hg (95% CI −2.45 to −1.01; *P*<.001), which closely aligned with the primary imputed analysis results (−1.86 mm Hg; *P*<.001). Similar consistency was observed for DBP and high-reading percentage outcomes. The convergence of estimates across missing data approaches, combined with the separation of imputed covariates from observed outcomes, provides strong evidence for the validity of our findings.

### Study 1: Association of the Digital Intervention With Improvements in High BP Levels

The results from the piecewise mixed-effects model revealed differences in the trajectories of the monthly average SBP between the intervention and control groups, as illustrated in Figure 5A. The intraclass correlation coefficient indicated that 62% of the total variance in SBP was attributable to between-subject differences.

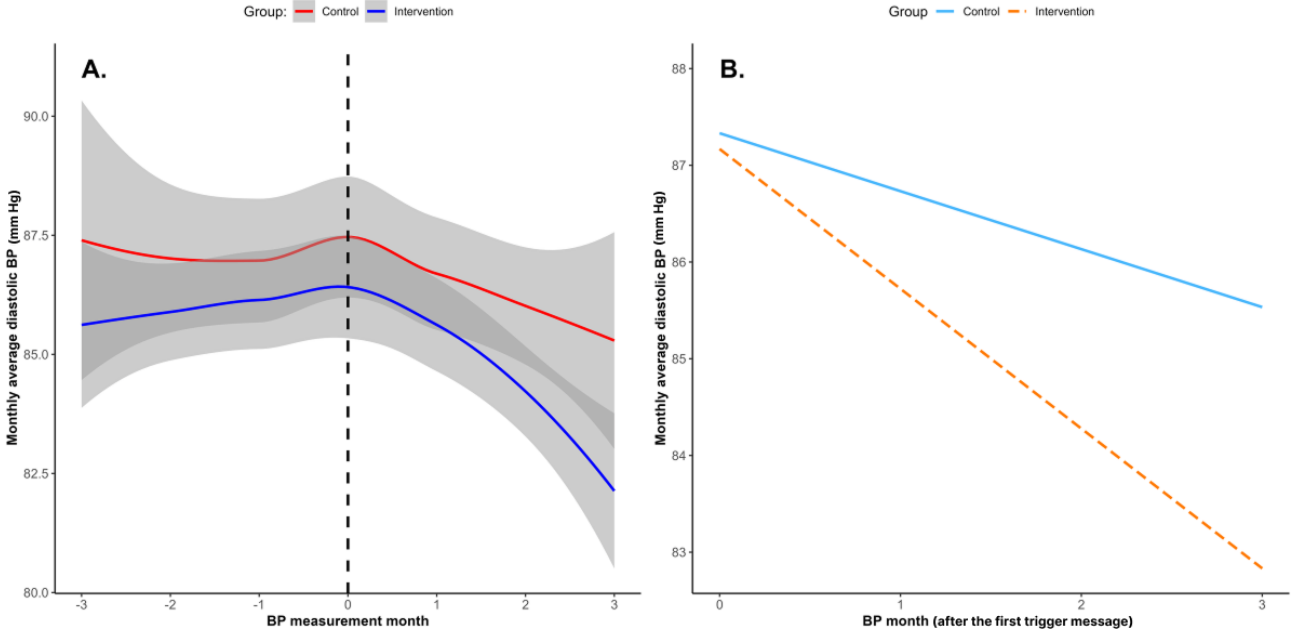
**Figure 5.** (A) Monthly average systolic blood pressure (BP) change over time. The dotted line represents the month when the first message was sent. (B) Simple slope analysis of the monthly average systolic BP during the period after the first nudge message was sent (moderated by group).




A significant interaction effect was observed following the delivery of the first nudge message (B=−1.04; *P*=.04), while no significant interaction was detected in the preceding period (B=−0.21; *P*=.67). The intervention group (n=148) demonstrated a significant decrease in the monthly average SBP during the 3 months following the nudge message (−2.09 mm Hg/month, 95% CI −2.77 to −1.41; *P*<.001), with a total reduction of 6.27 mm Hg over the 3-month period. The control group (n=148) also exhibited a significant reduction in the monthly average SBP (−1.06 mm Hg/month, 95% CI −1.83 to −0.28; *P*=.007), with a total reduction of 3.18 mm Hg over 3 months. The differential effect between groups (1.04 mm Hg/month) represents the additional benefit attributable to the digital nudges (Figure 5B).

The results of the piecewise mixed-effects model highlighted differences in the trajectories of the monthly average DBP between the groups, as shown in Figure 6A. The intraclass correlation coefficient indicated that 66% of the total variance in DBP was attributable to between-subject differences.

**Figure 6.** (A) Monthly average diastolic blood pressure (BP) change over time. The dotted line represents the month when the first message was sent. (B) Simple slope analysis of the monthly average diastolic BP during the period after the first nudge message was sent (moderated by group).

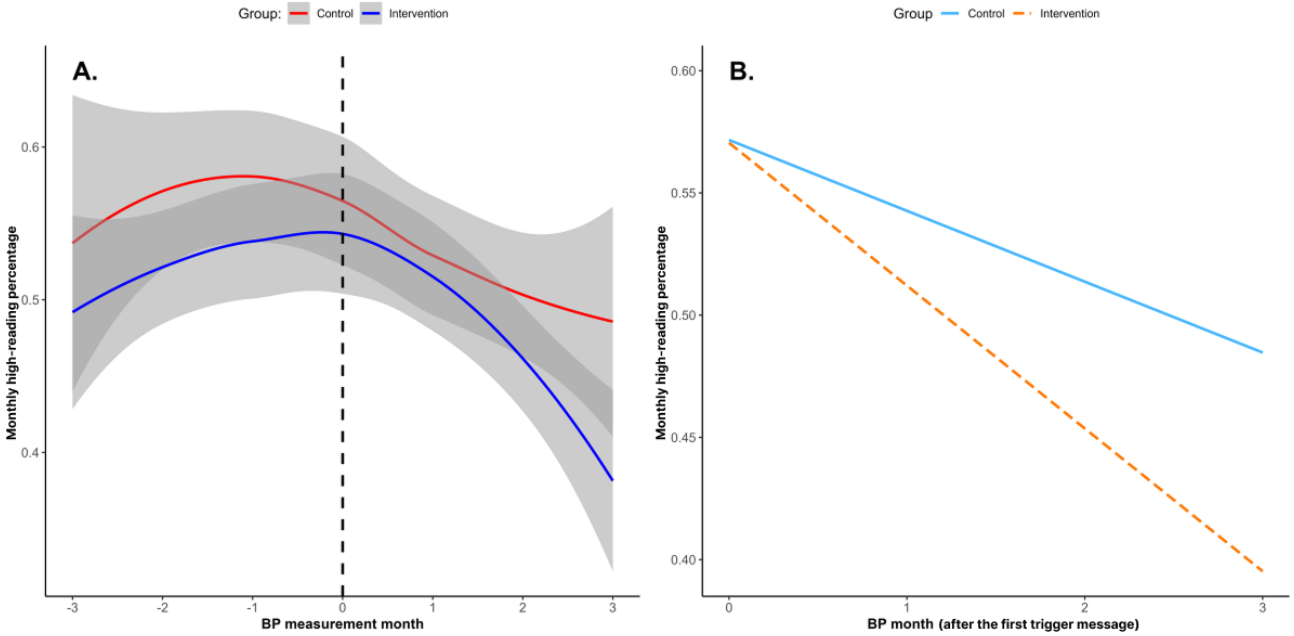



A significant interaction effect was observed during the period after the first nudge message was sent (B=−0.85; *P*=.02), while the interaction effect was not significant in the preceding period (B=−0.17; *P*=.61).

The intervention group exhibited a significant decrease in the monthly average DBP within the 3 months following the nudge message (−1.45 mm Hg/month, 95% CI −1.91 to −0.98; *P*<.001), with a total reduction of 4.35 mm Hg over the 3-month period. The control group also showed a significant reduction in the monthly average DBP (−0.60 mm Hg/month, 95% CI −1.13 to −0.07; *P*=.03), with a total reduction of 1.80 mm Hg over 3 months. The differential effect between groups (0.85 mm Hg/month) indicated that the intervention group experienced an additional 2.55 mm Hg reduction in DBP attributable to the digital nudges over the 3-month period (Figure 6B).

Another piecewise mixed-effects model was used to assess differences in the trajectories of the monthly average high-reading percentage between the groups, as illustrated in Figure 7A. The intraclass correlation coefficient indicated that 54% of the total variance in the monthly average high-reading percentage was attributable to between-subject differences.

**Figure 7.** (A) Monthly average high-reading percentage change over time. The dotted line represents the month when the first message was sent. (B) Simple slope analysis of the monthly average high-reading percentage during the period after the first nudge message was sent (moderated by group).



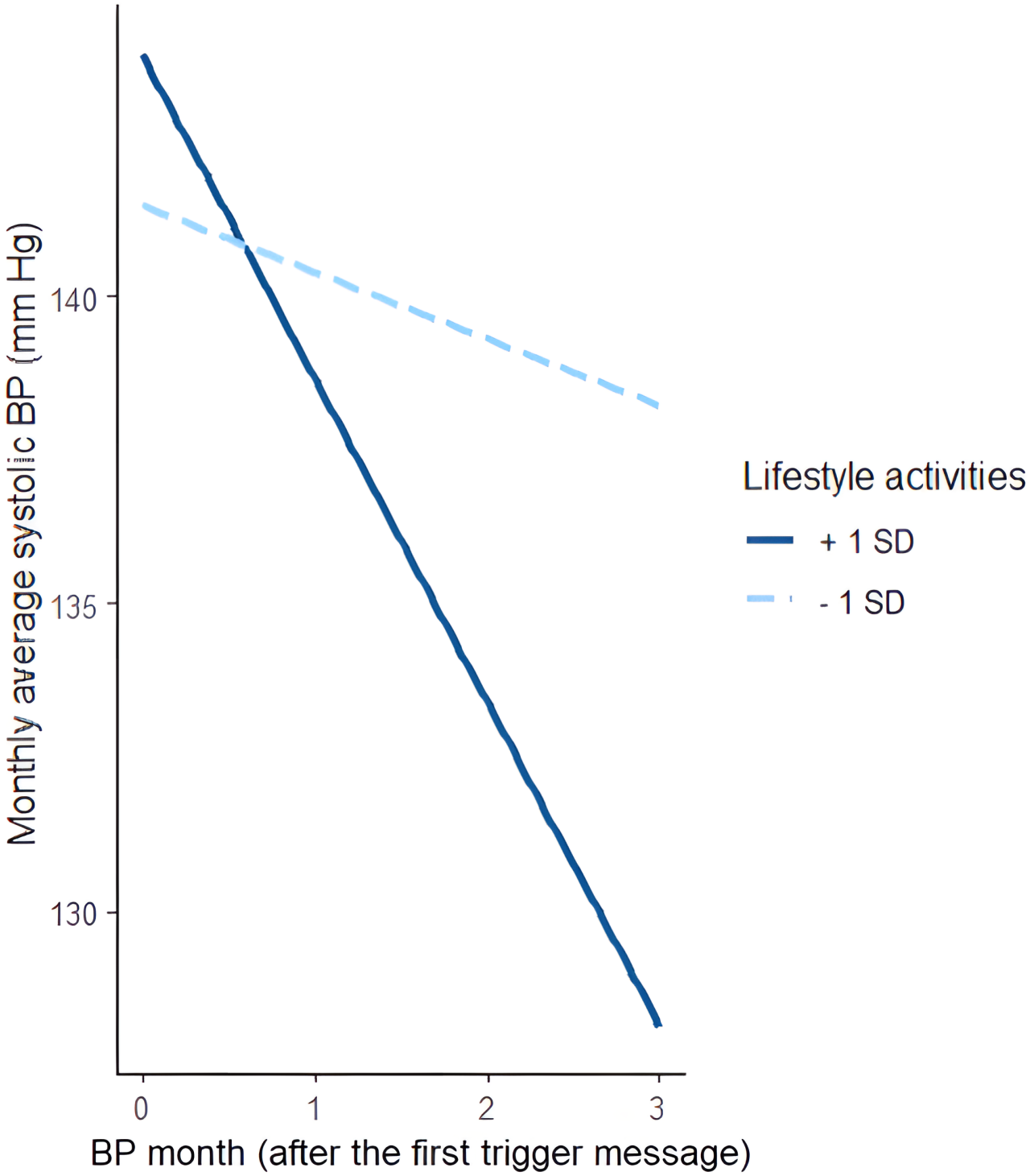

A significant interaction effect was observed during the period after the first nudge message was sent (B=−0.029; *P*=.04), while no significant interaction was found in the prior period (B=0.0005; *P*=.97). The intervention group exhibited a

significant decrease in the monthly high-reading percentage in the 3 months following the nudge message (−5.8 percentage points/month, 95% CI −7.7 to −4.0; *P*<.001), with a total reduction of 17.4 percentage points over the 3-month period. The control group also showed a significant reduction in the monthly high-reading percentage (−2.9 percentage points/month, 95% CI −5.0 to −0.8; *P*=.008), with a total reduction of 8.7 percentage points over 3 months. The differential effect between groups (2.9 percentage points/month) indicated that the intervention group experienced an additional 8.7 percentage point reduction in high readings attributable to the digital nudges over the 3-month period (Figure 7B).

Finally, a piecewise mixed-effects model was applied exclusively to the intervention group to evaluate the interaction between the 2 time periods and lifestyle activities in terms of the monthly average SBP. The results revealed a significant interaction effect between the number of lifestyle activities (operationalized as the sum of meal logs, carbohydrate intake, calories burned, and recipe finder activities in each month) and the monthly average SBP during the postmessage period (B=−0.12; *P*=.004), whereas no significant effect was observed in the premessage period (B=0.13; *P*=.33).

A simple slope analysis indicated that users with a higher monthly number of lifestyle activities experienced significant reductions in the monthly average SBP (B=−5.27; *P*<.001), while users with a below-average monthly number of lifestyle activities showed no significant changes (B=−1.08; *P*=.23) (Figure 8).

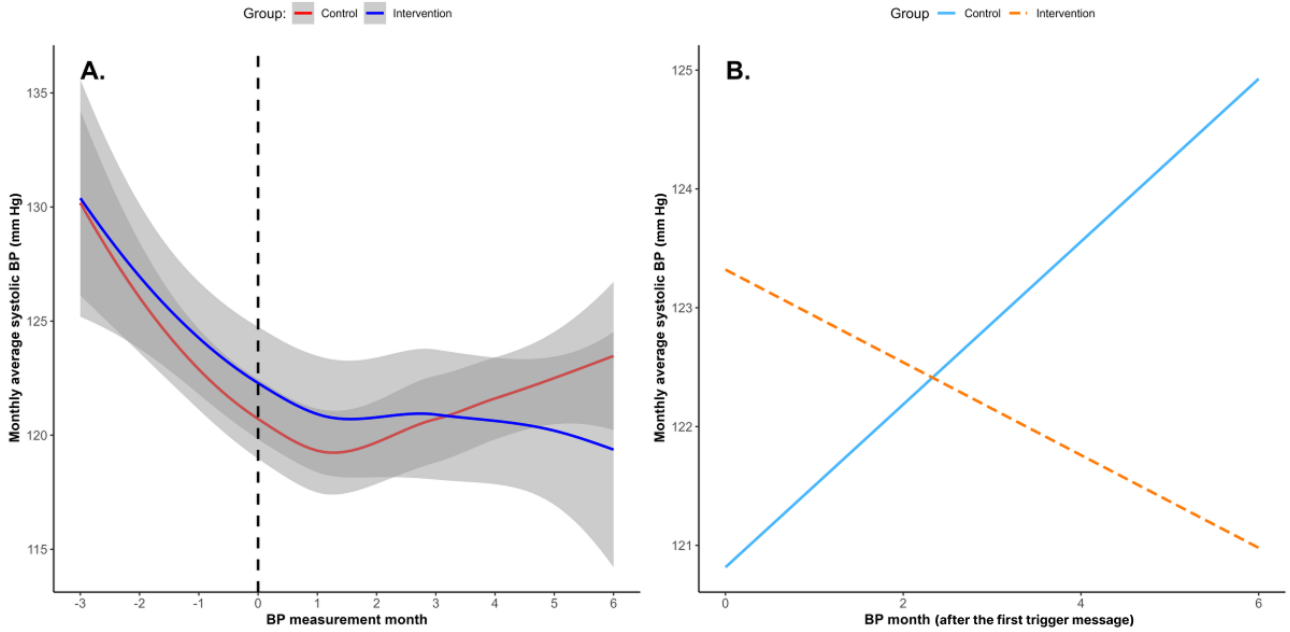
**Figure 8.** Simple slope analysis of the monthly average systolic blood pressure (BP) during the period after the first nudge message was sent (moderated by the number of lifestyle activities).




### Study 2: Association of the Digital Intervention With the Sustainability of Normal BP Levels

The results of the piecewise mixed-effects model highlighted differences in the trajectories of the monthly average SBP between the groups, as shown in Figure 9A. The intraclass correlation coefficient indicated that 36% of the total variance in SBP was attributable to between-subject differences.

**Figure 9.** (A) Monthly average systolic blood pressure (BP) change over time. The dotted line represents the month when the first message was sent. (B) Simple slope analysis of the monthly average systolic BP during the period after the first nudge message was sent (moderated by group).

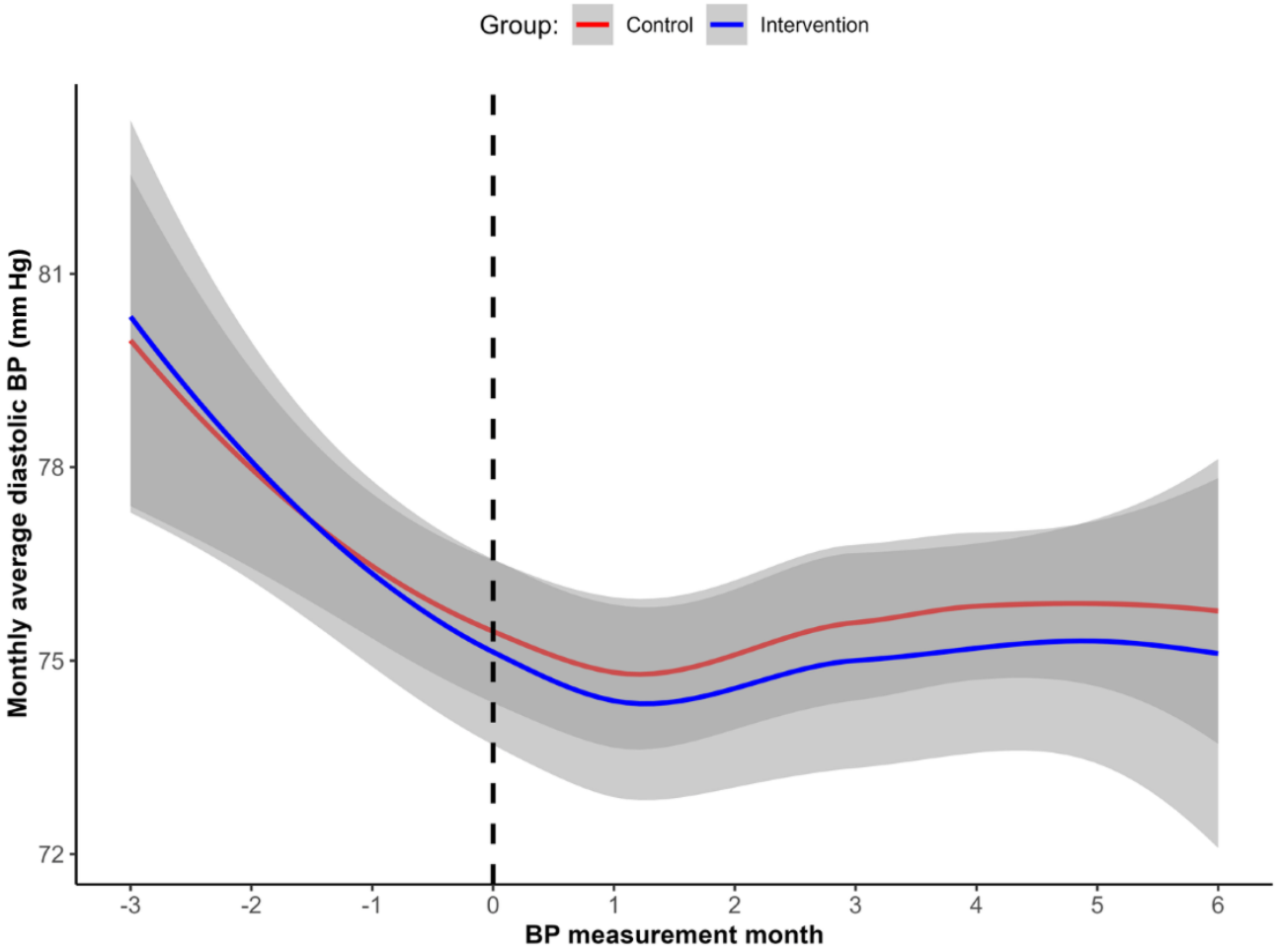



A significant interaction effect was observed in the 6-month period following the first nudge message (B=−1.08; *P*=.02), whereas no significant interaction effect was found in the period prior to the message (B=1.10; *P*=.32).

A simple slope analysis revealed that in the months following the messages, the intervention group (n=56) sustained stable BP levels with no significant change in the monthly average SBP (−0.39 mm Hg/month, 95% CI −1.08 to 0.30; *P*=.27). In contrast, the control group (n=56) exhibited a significant increase in the monthly average SBP (0.69 mm Hg/month, 95% CI 0.08 to 1.29; *P*=.03), with a total increase of 4.14 mm Hg over the 6-month period. The differential effect between groups (1.08 mm Hg/month) indicated that the digital nudges prevented an increase of approximately 6.48 mm Hg in SBP over the 6-month follow-up period compared to the control group (Figure 9B).

Another piecewise linear mixed-effects model was used to examine the differences in the trajectories of the monthly average DBP between the groups over the 3 months preceding and the 6 months following the first nudge message (Figure 10). The intraclass correlation coefficient indicated that 38% of the total variance in DBP was attributable to between-subject differences. The results revealed that neither group showed significant changes in DBP during the 6-month period following the intervention. The interaction effects between groups were not significant either before or after the messages (before the messages: B=−0.31; *P*=.64; after the messages: B=−0.13; *P*=.63), indicating no differential effect of the digital nudges on BP in the normal BP cohort. Both the intervention and control groups maintained stable DBP levels throughout the study period (Figure 10).

**Figure 10.** Monthly average diastolic blood pressure (BP) change over time. The dotted line represents the month when the first message was sent.




## Discussion

### Principal Findings

This retrospective study evaluated the effectiveness of a digital intervention aimed at helping individuals with high BP to regulate their levels by providing tailored digital support at the right time. This support was customized to individual needs and was compared to a matched control group that did not receive it. The findings revealed that users with higher engagement in lifestyle activities experienced significant reductions in their monthly average SBP. Additionally, the study suggested the critical role of positive support through messaging for maintaining normal BP levels. This support may enhance an individual's overall well-being and encourage positive behaviors that help sustain healthy BP levels over time and thus may support long-term hypertension management.

Our study used propensity score matching for the control group and used a piecewise mixed model as a statistical framework to describe the nonlinear behavior in BP levels, comparing 2 groups over time.

In the high BP cohort, our analysis indicated that before the intervention phase, the high readings in both groups demonstrated similar trajectories in terms of BP levels. However, after the start of the nudges, the BP levels in the intervention group significantly reduced, while the BP levels in the control group experienced a less pronounced effect and showed only a slight or less significant change. In the normal BP cohort, no significant interaction effect was found in the period prior to the positive feedback messages, while after the positive feedback messages were delivered, the intervention group sustained the reduction in their monthly average SBP, while the control group exhibited a significant increase in their monthly average SBP in the following 6-month period.

Robust statistically significant effects were found for improved diet, aerobic exercise, alcohol and sodium restriction, and fish oil supplements (mean reductions in SBP of 5.0, 4.6, 3.8, 3.6, and 2.3 mm Hg, respectively), with corresponding reductions in DBP [46].

Hypertension is a major cause of cardiovascular disease and kidney disease, and BP control reduces the risk of these complications [47,48]. Despite the availability of effective hypertensive medications, BP treatment and control rates remain low [48]. Self-management education, including user education, self-monitoring of clinical measurements, lifestyle modifications (such as a healthy diet, physical activity, weight management, smoking cessation, and alcohol reduction), and support for medication adherence, has been widely used for BP management [49,50]. Providing information about hypertension and its treatment, along with home BP monitoring and feedback, has been incorporated into effective interventions, demonstrating that supported self-management can improve BP control [50].

Healthy lifestyle and diet are associated with significant reductions in the risks of obesity, type 2 diabetes, high BP, and cardiovascular diseases [51]. Diet plays a crucial role in BP regulation through multiple mechanisms such as improved endothelial function, enhanced pressure natriuresis, and reduced oxidative stress [52]. Dietary management is considered to be an effective treatment method for hypertensive users as diet has a direct link with BP regulation [53,54]. Previous studies have suggested that lowering sodium intake mitigates hypertension symptoms significantly by lowering BP [55]. Conversely, increased sodium intake activates the hormonal mechanisms implicated in the pathogenesis of hypertension [55].

Recent studies have identified potential mechanisms by which physical activity enhances vascular function [56]. Regular exercise enhances vascular health through improvements in endothelial function. Specifically, exercise training in individuals with stable coronary artery disease has been shown to improve agonist-mediated, endothelium-dependent vasodilatory capacity [56]. Aerobic and resistance exercises have physiologically meaningful effects on endothelial function [57], increasing both the gene and protein expressions of endothelial nitric oxide synthase and boosting nitric oxide production in patients with coronary artery disease [51,58], which can result in arterial vasodilation and reduced peripheral resistance. This mechanism directly contributes to lower BP [59].

This study demonstrated that the use of automated data-driven nudges for interventions in hypertension can significantly influence BP levels over time. Digital messages may be effective for BP control by promoting lifestyle changes and improving medication adherence. Such interventions can encourage users to not only adopt healthier habits but also stay adherent to prescribed treatments or seek medical consultation for therapy adjustments in different populations, including low- and middle-income populations [60].

One of the broad mechanisms underlying the benefits of digital health tools is biofeedback to improve monitoring and management [24]. Digital health nudges impact the implementation and adherence to guideline-driven, evidence-based nonpharmacological strategies for reducing BP. This includes raising the awareness of how to properly measure BP at home, understanding BP categories and risk levels, promoting a healthy diet, reducing dietary salt intake, encouraging physical activity, managing stress, improving sleep hygiene, and addressing smoking and alcohol consumption [24,29,33,46,61]. The different components may also interact. For example, both stress reduction and decreased salt intake could contribute to better sleep quality [62].

The DASH diet has been reported to significantly lower SBP at every sodium level and significantly lower DBP at the high and intermediate sodium levels [33]. Through feedback messages, users receive guidance on proper medication use, including tips to improve adherence. Adherence is important, and high adherence to antihypertensive medication is associated with higher odds of BP control. Nonadherence to cardioprotective medication increases a user's risk of death from 50% to 80% [63]. Multiple factors contribute to adherence levels, stemming from individual, provider, and health care system elements, which often interact with each other [64]. Developments in the field of digital adherence monitoring and management are rapidly evolving to different cutting-edge solutions ranging from smartphone apps to smart devices. Technical and social innovations could lead to improved medication adherence and better user outcomes [34,65]. Self-monitoring of BP using a digital home BP monitoring device is another essential component of digital health in users with hypertension [23,29]. Use of out-of-office BP monitoring is preferred over office BP monitoring for the diagnosis and management of hypertension in major hypertension guidelines [29].

In the intervention presented in this manuscript, decisions concerning when and how to provide support are intervention-determined rather than user-determined. Tailoring variables were determined through active assessment of BP levels, which required user engagement. Users included in the study received prompts about their BP status when high readings (elevated and stage 1 or stage 2) were detected on 3 separate days within a 7-day period.

Real-time behavioral support that directly addresses user needs, where the content or timing of support is adapted and based on system-collected input and is triggered by the system, has been shown to be effective in digital health interventions [25]. The use of mHealth technology has been recommended as portable devices make intervention delivery more interactive and responsive by sending feedback messages in real time based on the user's state (eg, to promote physical activity and reduce sedentary behavior) [66,67]. Progress in digital health technology has enabled the design of interventions that aim to deliver behavior change support in real time, and this is matched to when users most want or need an intervention or when they are at risk. One of the common behavior change techniques is prompts/cues or feedback on behavior and action planning [25,68]. Digital health interventions, such as remote BP monitoring and tailored feedback, have been shown to be associated with significant reductions in BP levels among diverse populations [69]. Importantly, the findings of this study align with the results of previous meta-analyses, showing that individuals who received digital nudges experienced a greater reduction in SBP compared to those in the control group [22,69-71]. One of the most successful mHealth interventions combined the features of tailored messages, interactive communication, and multifaceted functions [71]. The findings of our study, which compared an intervention group to a control group using the same platform with or without the intervention, highlight the benefits of data-driven nudges that deliver real-time messages. The results suggest that adaptive, system-triggered support can effectively assist individuals in managing their BP. These outcomes could inform the development of an artificial intelligence–driven platform, which could act as a precise lifestyle guide for individuals with hypertension. Such a system could integrate BP monitoring with lifestyle data and leverage personalized machine learning models to assess the individual impact of various factors on BP. Previous programs typically provided users with remote monitoring devices and paired them with health coaches, but did not account for the individual

impact of lifestyle factors on BP, which can vary due to physiological differences [72]. In addition, physicians are often unable to optimally counsel patients on lifestyle modifications or personalize their guidance due to time constraints related to workload [73,74]. BP and lifestyle data collected from digital health platforms may allow us to view trends and make personalized recommendations to users. Personalized recommendations were found to be more useful compared to generic recommendations [72]. We believe that the combination of personalized guidance, ease of adherence, and motivational reinforcement contributed to high engagement and improved BP outcomes.

Our study also sheds light on the effect of positive feedback on normal BP levels in the long term, indicating that integrating positive behavioral principles can help users build resilience, improve motivation, and enhance user engagement. Positive feedback in digital health nudges has been shown to enhance user engagement and lead to better clinical outcomes in BP management through BP monitoring with supportive digital applications, which can increase user awareness and facilitate behavioral change, resulting in improved BP control [23].

Previous studies highlighted the importance of user empowerment tools, such as user-reported outcome measures and shared decision-making processes, in managing chronic diseases [75]. The integration of positive feedback within digital health interventions plays a crucial role in motivating users to adhere to their management plans, as demonstrated in previous studies [76]. We believe that incorporating positive feedback on normal BP levels enhances user motivation, fosters a hopeful mindset, and reinforces user confidence in making improvements. Additionally, recognizing and leveraging users' strengths can boost self-efficacy and encourage active engagement in their treatment plans [77]. In contrast, users who do not receive positive reinforcement may be less aware, potentially leading to increased BP levels over time.

By incorporating these positive psychology messages into clinical care, health care providers can support users in a holistic way, improving not only clinical outcomes but also users' quality of life and overall health management.

Additional research is needed to explore the best way to integrate user-determined features into a system-driven intervention, balancing structured, externally initiated support with individual autonomy in managing hypertension and other metabolic chronic conditions [78].

## Limitations

This study had several limitations. As with all retrospective real-world data analyses, the groups were not randomly assigned and treatment protocols were not prescribed, creating challenges in drawing causal inferences. Statistical modeling can address some difficulties in comparing groups and allow quasi-causal inferences; however, unmeasured variables may impact group balance.

While we carefully selected covariates for propensity score matching based on clinical relevance and prior literature, unmeasured confounding remains possible. Factors, including socioeconomic status, health literacy, dietary habits, stress levels, and intrinsic motivation for health behavior change, were not directly measured and could influence outcomes. Critically, this study lacked information on concurrent antihypertensive medication use and changes, which could significantly impact BP trajectories. Several aspects of our design helped mitigate these concerns. Matching on baseline BP measurement frequency partially captured health engagement, and the balance achieved on measured covariates (Figures 3 and 4) reduced the likelihood of substantial unmeasured confounding. Moreover, as the intervention and control groups experienced identical BP event patterns, unmeasured factors likely affected both groups similarly. Nevertheless, unmeasured characteristics that were differentially distributed between the groups could have influenced our effect estimates, an inherent limitation of a retrospective observational design. The COVID-19 pandemic was a major uncontrolled event separating the control and intervention periods, introducing a potential secular trend bias. Broader societal changes, such as increased health awareness and health care access, may have influenced user behavior and engagement patterns independent of the digital nudges.

The study population was limited to individuals who self-selected the use of the Dario system, indicating pre-existing health care engagement. Users who opened and read intervention messages may have been particularly motivated to change. However, our inclusion criteria ensured that both groups showed evidence of engagement with hypertension management, with no significant differences in BP measurement frequency between the groups, suggesting that motivation may not be the primary differentiating factor.

We acknowledge that external validity may be limited to health-engaged, technologically enabled populations, and the results may not be generalizable to those with limited access or comfort with digital tools. However, the growing use of smartphones and digital health across diverse groups suggests cautious optimism for broader applicability. Future research should focus on strategies to engage lower-motivation users and assess effectiveness in more diverse, underserved populations.

Additionally, Bluetooth-connected devices operating outside electronic medical records face adoption challenges, particularly for users with limited health technology literacy. The successful device adoption in our study population may limit generalizability. Nevertheless, the clinical benefits observed align with growing evidence that digital health interventions can achieve meaningful outcomes even without full electronic medical record integration, particularly when combined with behavioral support features like our automated nudge system.

Our temporal analysis design focused on monthly intervals over a 3-month (high BP cohort) or 6-month (normal BP cohort) period before and after the intervention. While relationships could potentially be investigated at daily or weekly scales, tracking such granular changes is difficult in real-world studies. Monthly average BP provides robust estimates of sustained intervention effects and aligns with clinical practice guidelines while providing sufficient statistical power to detect clinically meaningful effects. However, this approach cannot capture short-term variability patterns, including intraday fluctuations,

morning surges, or circadian patterns that have independent prognostic values for cardiovascular outcomes. Future studies with higher-frequency measurements could investigate whether digital interventions affect these BP variability parameters.

Critically, we cannot adjust for time-varying confounders over the 6-month study period. Our piecewise mixed-effects model accounts for individual trajectories, but it cannot control postbaseline medication changes, clinical visits, or seasonal BP variations. The use of historical controls versus intervention recipients can introduce a potential secular trend bias. An additional limitation of this study is the lack of stratified analysis on the intensity or frequency of digital nudges and their specific effects on BP outcomes. While the intervention demonstrated overall effectiveness, we did not assess whether higher or lower exposure to messaging would yield differential impacts on SBP or DBP outcomes. Future research should investigate the differential effectiveness of nudge intensity and timing, using experimental or adaptive designs, which will lead to more precise and effective hypertension management at scale.

Finally, available demographic and medical data were limited. Although no differences existed between the groups in age, gender, median household income, diabetes type, weight, or comorbidities, uncontrolled bias from other demographic or medical factors remains a possibility. Future studies should incorporate comprehensive medication data to better isolate digital intervention effects and explore strategies to reduce adoption barriers while maintaining demonstrated clinical effectiveness.

## Conclusions

This study provides evidence suggesting that digital health data–driven nudges may help support BP management by delivering real-time, adaptive behavioral support. Users who engaged more actively with lifestyle interventions experienced greater reductions in SBP, indicating a potential association between engagement and improved outcomes. Our findings highlight the significance of integrating positive messaging into digital health solutions. Positive feedback on normal BP levels may have encouraged users to maintain healthy behaviors, reinforcing motivation and self-efficacy. In contrast, those who did not receive positive feedback exhibited an upward trend in BP levels over time. This underscores the value of continuous engagement and reinforcement strategies in sustaining long-term hypertension management.

The use of a digital health platform enabled real-time monitoring, while data-driven nudges further enhanced user engagement by prompting timely self-monitoring, reinforcing healthy behaviors, and supporting adherence to lifestyle modifications and antihypertensive therapy. The application of machine learning models and artificial intelligence–driven platforms could further enhance the personalization of digital health interventions, tailoring recommendations based on individual responses to behavioral and physiological data.

## Future Research Directions

Despite the promising results, additional research is required to refine digital health interventions, ensuring they are both system-driven and adaptable to user preferences. Future studies should explore the optimal integration of user-determined features within structured, externally initiated interventions. Additionally, expanding these approaches to other cardiometabolic conditions could further validate the effectiveness of digital health tools in managing chronic diseases. By leveraging personalized recommendations, behavioral reinforcement, and positive feedback, digital health platforms have the potential to transform hypertension care, empowering individuals to take an active role in their health and achieve sustainable improvements in BP control.

## Acknowledgments

Generative artificial intelligence was used to improve the language and readability of the manuscript. This research was funded by DarioHealth. No external financial support or grants were received for the research, authorship, or publication of this article.

## Data Availability

The datasets generated or analyzed during this study are not publicly available owing to the privacy policy of the company but are available from the corresponding author upon reasonable request, subject to company policies.

## Authors' Contributions

Conceptualization: MDR, DLH, YF-H
Data curation: IBA
Formal analysis: PG, IBA
Investigation: IBA, YF-H
Methodology: PG, IBA, YF-H
Resources: YF-H
Supervision: YF-H
Validation: PG
Visualization: CL
Writing – original draft: IBA, CL, YF-H
Writing – review & editing: PG, MDR, DLH, OM

## Conflicts of Interest

IBA, CL, OM, and YF-H are employees of DarioHealth. MDR and DLH serve as scientific advisory board members of DarioHealth. PG has received a consulting fee to assist with the analysis but otherwise has no conflicts of interest.

## References


1. Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. Lancet 2005;365(9455):217-223. [doi: 10.1016/S0140-6736(05)17741-1] [Medline: 15652604]
2. GBD 2013 Risk Factors Collaborators, Forouzanfar MH, Alexander L, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Dec 5;386(10010):2287-2323. [doi: 10.1016/S0140-6736(15)00128-2] [Medline: 26364544]
3. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. Nat Rev Nephrol 2020 Apr;16(4):223-237. [doi: 10.1038/s41581-019-0244-2] [Medline: 32024986]
4. Hare AJ, Chokshi N, Adusumalli S. Novel digital technologies for blood pressure monitoring and hypertension management. Curr Cardiovasc Risk Rep 2021;15(8):11. [doi: 10.1007/s12170-021-00672-w] [Medline: 34127936]
5. Ostchega Y, Fryar CD, Nwankwo T, Nguyen DT. Hypertension prevalence among adults aged 18 and over: United States, 2017-2018. NCHS Data Brief 2020 Apr(364):1-8. [Medline: 32487290]
6. Hypertension - Health, United States. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/hus/topics/hypertension.htm [accessed 2024-12-16]
7. Rana J, Oldroyd J, Islam MM, Tarazona-Meza CE, Islam RM. Prevalence of hypertension and controlled hypertension among United States adults: evidence from NHANES 2017-18 survey. Int J Cardiol Hypertens 2020 Dec;7:100061. [doi: 10.1016/j.ijchy.2020.100061] [Medline: 33447782]
8. Farhadi F, Aliyari R, Ebrahimi H, Hashemi H, Emamian MH, Fotouhi A. Prevalence of uncontrolled hypertension and its associated factors in 50-74 years old Iranian adults: a population-based study. BMC Cardiovasc Disord 2023 Jun 24;23(1):318. [doi: 10.1186/s12872-023-03357-x] [Medline: 37355590]
9. Abrahamowicz AA, Ebinger J, Whelton SP, Commodore-Mensah Y, Yang E. Racial and ethnic disparities in hypertension: barriers and opportunities to improve blood pressure control. Curr Cardiol Rep 2023 Jan;25(1):17-27. [doi: 10.1007/s11886-022-01826-x] [Medline: 36622491]
10. Yang JH, Ham CH, Hwang YG, Suh SW. Metallosis: a complication in the guided growing rod system used in treatment of scoliosis. Indian J Orthop 2017;51(6):714-718. [doi: 10.4103/0019-5413.217692] [Medline: 29200491]
11. Neumann FJ, Sousa-Uva M, Ahlsson A, et al. 2018 ESC/EACTS guidelines on myocardial revascularization. Russ J Cardiol 2018 Jul;40(8):151-226. [doi: 10.15829/1560-4071-2019-8-151-226]
12. Parati G. Blood pressure variability: its measurement and significance in hypertension. J Hypertens 2005;23(Suppl 1):S19-S25. [doi: 10.1097/01.hjh.0000165624.79933.d3]
13. Mills KT, Bundy JD, Kelly TN, et al. Global disparities of hypertension prevalence and control: a systematic analysis of population-based studies from 90 countries. Circulation 2016 Aug 9;134(6):441-450. [doi: 10.1161/CIRCULATIONAHA.115.018912] [Medline: 27502908]
14. Carey RM, Muntner P, Bosworth HB, Whelton PK. Prevention and control of hypertension: JACC health promotion series. J Am Coll Cardiol 2018 Sep 11;72(11):1278-1293. [doi: 10.1016/j.jacc.2018.07.008] [Medline: 30190007]
15. Ogedegbe G. Barriers to optimal hypertension control. J of Clinical Hypertension 2008 Aug;10(8):644-646. [doi: 10.1111/j.1751-7176.2008.08329.x]
16. Oparil S, Acelajado MC, Bakris GL, et al. Hypertension. Nat Rev Dis Primers 2018 Mar 22;4(1):18014. [doi: 10.1038/nrdp.2018.14] [Medline: 29565029]
17. Liu L, Guo J, Jiao X, You L. The effect of digital therapeutics intervention on improving hypertension management in adults: a meta-analysis of randomized controlled trial. Hypertens Res 2025 Feb;48(2):456-469. [doi: 10.1038/s41440-024-01892-4] [Medline: 39443703]
18. Matthes J, Albus C. Improving adherence with medication. Deutsches Ärzteblatt International 2014 Jan;111(4). [doi: 10.3238/arztebl.2014.0041] [Medline: 24612495]
19. Uchmanowicz B, Chudiak A, Uchmanowicz I, Rosińczuk J, Froelicher ES. Factors influencing adherence to treatment in older adults with hypertension. Clin Interv Aging 2018;13:2425-2441. [doi: 10.2147/CIA.S182881] [Medline: 30568434]
20. David C, Ziegelmann PK, Goveia P, Silvani J, Silveira L, Fuchs SC. The effect of mobile health focused on diet and lifestyle on blood pressure: a systematic review and meta-analysis. Eur J Prev Cardiol 2022 May 25;29(7):1142-1155. [doi: 10.1093/eurjpc/zwac017]
21. Siopis G, Moschonis G, Eweka E, et al. Effectiveness, reach, uptake, and feasibility of digital health interventions for adults with hypertension: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health 2023 Mar;5(3):e144-e159. [doi: 10.1016/S2589-7500(23)00002-X] [Medline: 36828607]
22. Lu X, Yang H, Xia X, et al. Interactive mobile health intervention and blood pressure management in adults. Hypertension 2019 Sep;74(3):697-704. [doi: 10.1161/HYPERTENSIONAHA.119.13273] [Medline: 31327259]

23. Fundoiano-Hershcovitz Y, Bacher D, Ritholz MD, Horwitz DL, Manejwala O, Goldstein P. Blood pressure monitoring as a digital health tool for improving diabetes clinical outcomes: retrospective real-world study. J Med Internet Res 2022 Feb 8;24(2):e32923. [doi: 10.2196/32923] [Medline: 35133284]
24. Kario K, Harada N, Okura A. Digital therapeutics in hypertension: evidence and perspectives. Hypertension 2022 Oct;79(10):2148-2158. [doi: 10.1161/HYPERTENSIONAHA.122.19414] [Medline: 35726619]
25. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019 Apr 3;16(1):31. [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
26. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
27. Reule S, Drawz PE. Heart rate and blood pressure: any possible implications for management of hypertension? Curr Hypertens Rep 2012 Dec;14(6):478-484. [doi: 10.1007/s11906-012-0306-3] [Medline: 22972532]
28. The relationship between your heart rate and blood pressure. Cleveland Clinic. URL: https://health.clevelandclinic.org/heart-rate-blood-pressure [accessed 2025-09-07]
29. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2018 Jun;71(6):e13-e115. [doi: 10.1161/HYP.0000000000000065] [Medline: 29133356]
30. Writing Committee Members, Jones DW, Ferdinand KC, Taler SJ, et al. 2025 AHA/ACC/AANP/AAPA/ABC/ACCP/ACPM/AGS/AMA/ASPC/NMA/PCNA/SGIM guideline for the prevention, detection, evaluation and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation 2025 Sep 16;152(11):e114-e218. [doi: 10.1161/CIR.0000000000001356] [Medline: 40811497]
31. Rubenfire M. 2017 guideline for high blood pressure in adults. American College of Cardiology. URL: https://www.acc.org/Latest-in-Cardiology/ten-points-to-remember/2017/11/09/11/41/2017-Guideline-for-High-Blood-Pressure-in-Adults [accessed 2025-02-02]
32. Dietary guidelines for Americans. Food and Nutrition Service. URL: https://www.fns.usda.gov/cnpp/dietary-guidelines-americans [accessed 2025-02-02]
33. Sacks FM, Svetkey LP, Vollmer WM, et al. Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet. DASH-Sodium Collaborative Research Group. N Engl J Med 2001 Jan 4;344(1):3-10. [doi: 10.1056/NEJM200101043440101] [Medline: 11136953]
34. Kardas P. From non-adherence to adherence: can innovative solutions resolve a longstanding problem? Eur J Intern Med 2024 Jan;119(6–12):6-12. [doi: 10.1016/j.ejim.2023.10.012] [Medline: 37848351]
35. Jimmy B, Jose J. Patient medication adherence: measures in daily practice. Oman Med J 2011 May;26(3):155-159. [doi: 10.5001/omj.2011.38] [Medline: 22043406]
36. Lally P, van Jaarsveld CHM, Potts HWW, Wardle J. How are habits formed: modelling habit formation in the real world. Euro J Social Psych 2010 Oct;40(6):998-1009. [doi: 10.1002/ejsp.674]
37. Ethical and Independent Review Services. URL: https://versiticlinicaltrials.org/about-us/our-brands/salus-irb [accessed 2020-08-23]
38. Terms of use. DarioHealth. URL: https://www.dariohealth.com/terms-of-use/ [accessed 2024-04-18]
39. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011 May;46(3):399-424. [doi: 10.1080/00273171.2011.568786] [Medline: 21818162]
40. Little RJA. A test of missing completely at random for multivariate data with missing values. J Am Stat Assoc 1988 Dec;83(404):1198-1202. [doi: 10.1080/01621459.1988.10478722]
41. Stuart EA. Matching methods for causal inference: a review and a look forward. Statist Sci 2010 Feb;25(1):1-21. [doi: 10.1214/09-STS313]
42. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med 2009 Nov 10;28(25):3083-3107. [doi: 10.1002/sim.3697] [Medline: 19757444]
43. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat 2011;10(2):150-161. [doi: 10.1002/pst.433] [Medline: 20925139]
44. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. Am J Epidemiol 2006 Jun 15;163(12):1149-1156. [doi: 10.1093/aje/kwj149] [Medline: 16624967]
45. Kohli N, Peralta Y, Zopluoglu C, Davison ML. A note on estimating single-class piecewise mixed-effects models with unknown change points. Int J Behav Dev 2018 Sep;42(5):518-524. [doi: 10.1177/0165025418759237]
46. Dickinson HO, Mason JM, Nicolson DJ, et al. Lifestyle interventions to reduce raised blood pressure: a systematic review of randomized controlled trials. J Hypertens 2006 Feb;24(2):215-233. [doi: 10.1097/01.hjh.0000199800.72563.26] [Medline: 16508562]

47. Melgarejo JD, Maestre GE, Thijs L, et al. Prevalence, treatment, and control rates of conventional and ambulatory hypertension across 10 populations in 3 continents. Hypertension 2017 Jul;70(1):50-58. [doi: 10.1161/HYPERTENSIONAHA.117.09188] [Medline: 28483916]
48. Unger T, Borghi C, Charchar F, et al. 2020 International Society of Hypertension Global Hypertension practice guidelines. Hypertension 2020 Jun;75(6):1334-1357. [doi: 10.1161/HYPERTENSIONAHA.120.15026] [Medline: 32370572]
49. Tuglo LS. Comparison of adiposity anthropometric indices and their associations with visceral fat levels determined by bioelectrical impedance analysis among diabetic patients. Sci Rep 2022 Oct 24;12(1):17824. [doi: 10.1038/s41598-022-22848-z] [Medline: 36280694]
50. Shahaj O, Denneny D, Schwappach A, et al. Supporting self-management for people with hypertension: a meta-review of quantitative and qualitative systematic reviews. J Hypertens 2019 Feb;37(2):264-279. [doi: 10.1097/HJH.0000000000001867] [Medline: 30020240]
51. Man AWC, Li H, Xia N. Impact of lifestyles (diet and exercise) on vascular health: oxidative stress and endothelial function. Oxid Med Cell Longev 2020;2020:1496462. [doi: 10.1155/2020/1496462] [Medline: 33062134]
52. Altawili AA, Altawili M, Alwadai AM, et al. An exploration of dietary strategies for hypertension management: a narrative review. Cureus 2023 Dec;15(12):e50130. [doi: 10.7759/cureus.50130] [Medline: 38186513]
53. Appel LJ. The effects of dietary factors on blood pressure. Cardiol Clin 2017 May;35(2):197-212. [doi: 10.1016/j.ccl.2016.12.002] [Medline: 28411894]
54. Saleem K, Abdi G, Rabail R, Aadil RM. Dietary and lifestyle modifications for effective management of hypertension: a comprehensive review. Journal of Agriculture and Food Research 2025 Apr;20:101750. [doi: 10.1016/j.jafr.2025.101750]
55. Huang L, Trieu K, Yoshimura S, et al. Effect of dose and duration of reduction in dietary sodium on blood pressure levels: systematic review and meta-analysis of randomised trials. BMJ 2020 Feb 24;368:m315. [doi: 10.1136/bmj.m315] [Medline: 32094151]
56. Hambrecht R, Adams V, Erbs S, et al. Regular physical activity improves endothelial function in patients with coronary artery disease by increasing phosphorylation of endothelial nitric oxide synthase. Circulation 2003 Jul 1;107(25):3152-3158. [doi: 10.1161/01.CIR.0000074229.93804.5C] [Medline: 12810615]
57. Banks NF, Rogers EM, Stanhewicz AE, Whitaker KM, Jenkins NDM. Resistance exercise lowers blood pressure and improves vascular endothelial function in individuals with elevated blood pressure or stage-1 hypertension. Am J Physiol Heart Circ Physiol 2024 Jan 1;326(1):H256-H269. [doi: 10.1152/ajpheart.00386.2023] [Medline: 37975709]
58. Lee JH, Lee R, Hwang MH, Hamilton MT, Park Y. The effects of exercise on vascular endothelial function in type 2 diabetes: a systematic review and meta-analysis. Diabetol Metab Syndr 2018;10:15. [doi: 10.1186/s13098-018-0316-7] [Medline: 29541164]
59. Yang Y, Lv Q, Hou X, et al. Effects of different exercise modalities on blood pressure and endothelial function in prehypertension individuals: a systematic review and network meta-analysis. Front Cardiovasc Med 2025 Jun 25;12. [doi: 10.3389/fcvm.2025.1550435]
60. Boima V, Doku A, Agyekum F, Tuglo LS, Agyemang C. Effectiveness of digital health interventions on blood pressure control, lifestyle behaviours and adherence to medication in patients with hypertension in low-income and middle-income countries: a systematic review and meta-analysis of randomised controlled trials. EClinicalMedicine 2024 Mar;69:102432. [doi: 10.1016/j.eclinm.2024.102432] [Medline: 38333367]
61. Pereira A, Trombini R, Barbalho Y, et al. Strategies for effective communication in hypertension management: validation of messages from a mobile application to assist hypertensive older adults in adherence to treatment, nutrition and physical activity. Nutrients 2024 Dec 12;16(24):4284. [doi: 10.3390/nu16244284] [Medline: 39770906]
62. Chen TL, Chang SC, Hsieh HF, Huang CY, Chuang JH, Wang HH. Effects of mindfulness-based stress reduction on sleep quality and mental health for insomnia patients: a meta-analysis. J Psychosom Res 2020 Aug;135:110144. [doi: 10.1016/j.jpsychores.2020.110144] [Medline: 32590218]
63. Improving medication adherence among patients with hypertension. Million Hearts. URL: https://millionhearts.hhs.gov/data-reports/factsheets/adherence.html [accessed 2025-02-02]
64. Chaudri NA. Adherence to long-term therapies evidence for action. Ann Saudi Med 2004;24(3):221-222. [doi: 10.5144/0256-4947.2004.221]
65. van Boven JFM, Fonseca JA. Editorial: digital tools to measure and promote medication adherence. Front Med Technol 2021;3(751976):751976. [doi: 10.3389/fmedt.2021.751976] [Medline: 35047961]
66. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. Ann Behav Med 2017 Apr;51(2):226-239. [doi: 10.1007/s12160-016-9846-0] [Medline: 27757789]
67. Müller AM, Alley S, Schoeppe S, Vandelanotte C. The effectiveness of e-& mHealth interventions to promote physical activity and healthy diets in developing countries: a systematic review. Int J Behav Nutr Phys Act 2016 Oct 10;13(1):109. [doi: 10.1186/s12966-016-0434-2] [Medline: 27724911]
68. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]

69. Katz ME, Mszar R, Grimshaw AA, et al. Digital health interventions for hypertension management in US populations experiencing health disparities: a systematic review and meta-analysis. JAMA Netw Open 2024 Feb 5;7(2):e2356070. [doi: 10.1001/jamanetworkopen.2023.56070] [Medline: 38353950]
70. Cao W, Milks MW, Liu X, et al. mHealth interventions for self-management of hypertension: framework and systematic review on engagement, interactivity, and tailoring. JMIR Mhealth Uhealth 2022 Mar 2;10(3):e29415. [doi: 10.2196/29415] [Medline: 35234655]
71. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e17776. [doi: 10.2196/17776] [Medline: 32217503]
72. Leitner J, Chiang PH, Agnihotri P, Dey S. The effect of an AI-based, autonomous, digital health intervention using precise lifestyle guidance on blood pressure in adults with hypertension: single-arm nonrandomized trial. JMIR Cardio 2024 May 28;8(1):e51916. [doi: 10.2196/51916] [Medline: 38805253]
73. Bell RA, Kravitz RL. Physician counseling for hypertension: what do doctors really do? Patient Educ Couns 2008 Jul;72(1):115-121. [doi: 10.1016/j.pec.2008.01.021]
74. Bakris G, Ali W, Parati G. ACC/AHA versus ESC/ESH on hypertension guidelines: JACC guideline comparison. J Am Coll Cardiol 2019 Jun 18;73(23):3018-3026. [doi: 10.1016/j.jacc.2019.03.507] [Medline: 31196460]
75. Erturkmen GBL, Juul NK, Redondo IE, et al. Design, implementation and usability analysis of patient empowerment in ADLIFE project via patient reported outcome measures and shared decision making. BMC Med Inform Decis Mak 2024 Jun 28;24(1):185. [doi: 10.1186/s12911-024-02588-y] [Medline: 38943152]
76. Peeples M, (Disha) Maity B. Digital health in chronic care and self-management. In: Kiel JM, Kim GR, Ball MJ, editors. Healthcare Information Management Systems Health Informatics: Springer; 2022:209-222. [doi: 10.1007/978-3-031-07912-2_14]
77. Perkins A. Motivational interviewing: addressing patient ambivalence to facilitate change. MDatl. URL: https://mdatl.com/2024/10/motivational-interviewing-addressing-patient-ambivalence-to-facilitate-change/ [accessed 2025-02-09]
78. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clinical Psychology: Science and Practice 2017;24(1):27-45. [doi: 10.1037/h0101740]


## Abbreviations

**AHA:** American Heart Association
**BP:** blood pressure
**DASH:** Dietary Approaches to Stop Hypertension
**DBP:** diastolic blood pressure
**JITAI:** just-in-time adaptive intervention
**SBP:** systolic blood pressure
**SMD:** standardized mean difference


*Edited by A Coristine; submitted 21.04.25; peer-reviewed by E Spatz, Z Zhou; revised version received 28.10.25; accepted 28.10.25; published 26.11.25.*

*Please cite as:*
*Fundoiano-Hershcovitz Y, Breuer Asher I, D Ritholz M, L Horwitz D, Manejwala O, Levi C, Goldstein P*
*The Impact of Digital Intervention Messages Targeting Users With High Blood Pressure Events: Retrospective Real-World Study*
*JMIR Cardio 2025;9:e76275*
*URL: https://cardio.jmir.org/2025/1/e76275*
*doi:10.2196/76275*

# Outcomes of Team-Based Digital Monitoring of Patients With Multiple Chronic Conditions: Semiparametric Event Study

Ross Graham[1], MA, MS; Itzik Fadlon[2], PhD; Parag Agnihotri[3], MD; Christopher A Longhurst[4,5], MD, MS; Ming Tai-Seale[5,6], PhD, MPH

[1]Department of Sociology, University of California San Diego, La Jolla, CA, United States

[2]Department of Economics, University of California, San Diego, La Jolla, CA, United States

[3]Department of Medicine, University of California San Diego, La Jolla, CA, United States

[4]Department of Pediatrics, University of California San Diego, La Jolla, CA, United States

[5]Department of Bioinformatics, University of California San Diego Medical Center, La Jolla, CA, United States

[6]Department of Family Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, United States

**Corresponding Author:**
Ming Tai-Seale, PhD, MPH
Department of Bioinformatics, University of California San Diego Medical Center, La Jolla, CA, United States

## Abstract

**Background:** Remote patient monitoring (RPM) has emerged as an effective strategy for controlling hypertension by enabling patients to collect and transmit blood pressure (BP) data outside the clinic and supporting proactive care team interventions. While its benefits for hypertension management are well established, less is known about its effectiveness in patients with multiple chronic conditions (MCC), who experience higher morbidity, mortality, and costs.

**Objective:** This study aimed to evaluate the impact of an electronic health record (EHR)–integrated, team-based RPM program on patients with hypertension, alone or co-occurring with ischemic heart disease, type 2 diabetes, or both. This study aimed to determine whether referral to the program was associated with reductions in systolic blood pressure (SBP) across these patient groups.

**Methods:** We analyzed EHR data from patients referred by their primary care physicians to the University of California San Diego Health's Digital Health Program between October 2020 and July 2022. Eligible patients had hypertension, either alone or accompanied by at least 1 coexisting condition, such as ischemic heart disease or type 2 diabetes. Participants received a Bluetooth-enabled BP cuff and ongoing support from a multidisciplinary team, including nurse care managers and a pharmacist. A semiparametric event study design was used to estimate changes in SBP over 24 months, comparing prereferral and postreferral outcomes. To understand the program's impact, outcomes were analyzed for the full cohort of all referred patients and then scaled to reflect the average change in SBP among the program participants.

**Results:** Among patients who had been referred to the program, those with hypertension only experienced an average reduction of 9.70 (SE 0.80) mm Hg in SBP by the end of the analysis horizon of 1 year. Patients with hypertension and either diabetes or ischemic heart disease experienced a reduction of 6.61 (SE 1.12) mm Hg, and those with all 3 conditions experienced a reduction of 6.60 (SE 1.72) mm Hg. The average reductions in SBP among active participants were 16.83 mm Hg, 13.22 mm Hg, and 16.01 mm Hg, respectively.

**Conclusions:** A team-based, EHR-integrated RPM program was associated with clinically meaningful SBP reductions among patients with MCC. The program leveraged existing EHR workflows for referral and monitoring and provided technical and clinical support to patients. These findings suggest that EHR-integrated RPM services can achieve substantial improvements in BP in high-risk populations. As reimbursement for RPM expands, such models represent a promising strategy for addressing hypertension and the disproportionate burden of MCC at the population level.

***(JMIR Cardio 2025;9:e75170)*** doi:10.2196/75170

**KEYWORDS**

remote patient monitoring; telemedicine; hypertension; digital medicine; multiple chronic conditions

## Introduction

### Hypertension and Multiple Chronic Conditions

Hypertension is a widespread chronic disease affecting at least 33% of people aged 30 to 79 years worldwide [1] and approximately 48% of US adults [2], and contributes to poor cardiovascular, metabolic, mental health, and other chronic conditions. Approximately 27% of US adults have multiple chronic conditions (MCC) [3]. This patient population comprises up to 80% of Medicare costs [4], despite being the least studied among individuals with chronic diseases [5], and requires the most overall health care [6]. Common co-occurring chronic conditions include hypertension, type 2 diabetes, and ischemic heart disease. Among Medicare beneficiaries with MCC, diabetes and hypertension were the fourth most common combination (29.6% of MCC patients) in men aged <65 years and the fifth most common combination in men and women aged >65 years (28.6% and 27.5%, respectively) [7]. Ischemic heart disease and hypertension were the fifth most common combination in men aged <65 years (24.6%) and the third most common combination in men aged >65 years (39.0%) [7]. In men, the triad of ischemic heart disease, hypertension, and diabetes was the fourth most common [7]. Another study of US adults found that co-occurring diabetes and hypertension was the second most likely combination of MCC in all men and women aged >44 years [8]. Hypertension, ischemic heart disease, and diabetes share common pathophysiological processes. Their co-occurrence contributes to an increased risk of cardiovascular events and all-cause mortality [9,10].

### Remote Patient Monitoring

Remote patient monitoring (RPM) programs have shown positive effects in controlling chronic diseases such as hypertension and hyperlipidemia, proving feasible and effective in improving clinical outcomes in diverse patient populations [11-14]. These programs leverage technologies to empower patients and health care providers with regular, consistent collection of clinical data within everyday contexts for patients. They increase patient confidence and agency in managing chronic diseases and lower clinical exacerbations and hospitalizations, costs, and travel inconvenience for patients [15,16]. However, few health delivery organizations have scaled these solutions broadly, often citing the start-up investments involved and the need to pivot toward improved IT infrastructure [16], as well as complications that arise from new billing procedures and regulations around data privacy [17]. Some physicians and patients have expressed concern about the technological requirements of RPM, either with respect to patient capacity to use the technology consistently and effectively or that the technology itself might not be reliable over a sustained period [15]. However, reviews suggest that in practice, these concerns are often unfounded [17]. Even patients with English as their second language adapt well to RPM when provided with multilingual tools for assistance [18,19]. Generally, combining RPM with nurse or counselor follow-up and structured self-management interventions, such as patient education sessions, increases its effectiveness [20]. RPM is mostly used for cardiovascular health management [21] but is also effective in other domains, for example, remote glucose monitoring for diabetes [22].

### RPM in Patients With Hypertension and MCC

RPM of blood pressure (BP) for hypertension care shows evidence of sustainable decreases in patient systolic blood pressure (SBP) [12,23]. It reduces instances of white coat hypertension diagnosis by relying on real-world, at-home readings as opposed to office visits [24]. RPM effectively aids rural and medically underserved individuals and could play an important role in bridging the care gaps in chronically underserved regions [25], as well as racial minorities [26]. However, RPM has been previously shown to lead to stronger clinical improvements among racial, ethnic, and geographic groups who already experience better clinical outcomes (eg, White patients and those who live in affluent localities) due to lower RPM adoption by Black or Hispanic patients and those who live in less advantaged communities [27,28]. These effects are magnified when skills-based education of patients and community clinical providers generally serving medically underserved populations accompanies the technical rollout of RPM [29]. RPM has effectively helped control BP in postpartum women with hypertension [30].

RPM is effective in reducing the clinical burden of patients with MCC. Clinical trials and real-world observational and cohort studies show that RPM can lower mortality risk and SBP among patients with MCC who have heart disease and hypertension [24,31]. An RPM trial of patients with hypertension and diabetes observed a 9.1-point SBP reduction, bringing 51% of participants under 130/80 mm Hg within a year [32]; another observed an 11-point SBP reduction over 4 months [33]. A recent prospective observational study of patients with hypertension and high cholesterol, offering RPM alongside a patient support system, yielded a 9.7-point SBP reduction after 12 months [14].

Therefore, this study aimed to evaluate the impact of an electronic health record (EHR)-integrated, team-based RPM program on SBP among patients with hypertension, with a focus on patients with MCC.

## Methods

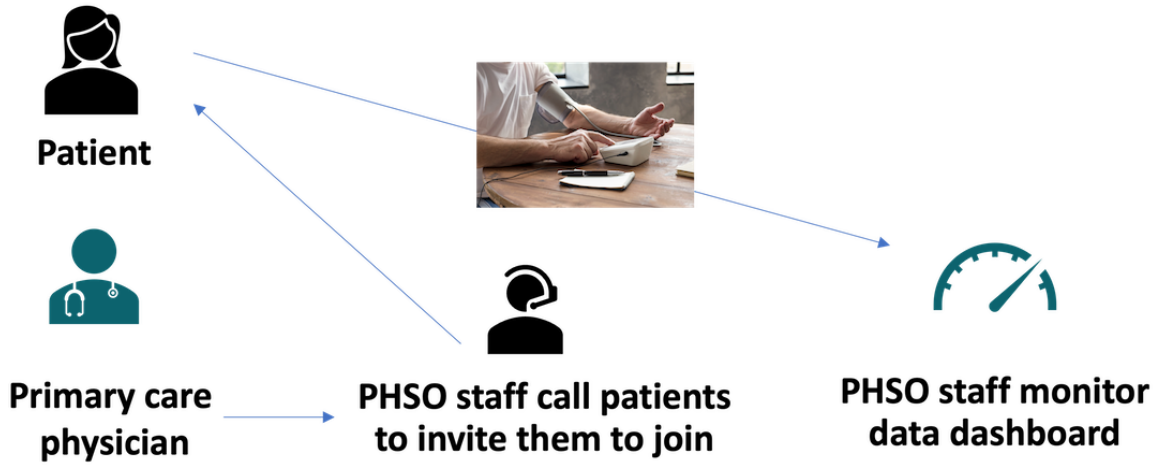
### The RPM Program

Hypertension management has been a priority at the University of California (UC) San Diego Health as part of a broader effort to mitigate cardiovascular risk among patients. The RPM service is a team-based, EHR-integrated service at the UC San Diego Health for patients with poorly managed hypertension. The program aims to lower cardiovascular risk while also reducing patient burden in managing their health and has been provided to a broad and diverse population of patients. Primary care physicians (PCPs) were introduced to the digital health service provided by the Population Health Services Organization via a presentation, either at departmental meetings or at a primary care retreat in 2021. They were shown how to refer eligible patients—those with a BP trend ≥140/90 mm Hg—within the EHR under a Care Coordination tab. PCPs were encouraged to refer patients as part of established practice guidelines as a

complement to their existing care practices. Once referred by their PCP, patients who were interested in participating completed a short questionnaire with the Digital Health team to confirm that their smartphone supported the app used to record BP measurements. This questionnaire ascertained that their digital literacy was sufficient to take daily BP readings by confirming that they were comfortable downloading and setting up a smartphone app and using Bluetooth. Eligible participating patients were then provided with iHealth Ease or Omron free of charge, both of which are Bluetooth-enabled digital BP meters that automatically transmit data to the patient portal of the EHR via a Health Insurance Portability and Accountability Act (HIPAA)-compliant smartphone app.

Enrollees are supported by a digital health specialist via telephone or, in cases where patients have extensive problems, via a visit to the patient's home for device setup and any technical issues. These measures attempt to mitigate equity issues arising from variations in technological literacy, which is crucial given that the target population tends to be older. The EHR system stratifies BP measures into 3 risk groups: normal, high or priority, and critical, and presents them on a daily dashboard. Nurse care managers and a pharmacist review the dashboard on a daily basis and make treat-to-target adjustments per protocol, encompassing medication adjustments and behavioral change recommendations. Participating patients were advised to take BP measurements daily and were contacted if they did not submit a reading at least once every 30 days. Patients were advised to participate in the program for at least 6 months and had technical support available for the entirety of their participation (Multimedia Appendix 1) [27]. This ongoing outreach incentivizes adherence, where integration of monitoring into digital technology minimizes the clinical burden while expanding patient inclusion and furthering equity. Outlined in Figure 1, the program workflow is initiated with the PCP referring to the patient via the EHR. A digital health specialist then calls the patient to onboard them, enabling the patient to transmit BP measurements that are then monitored by the Population Health Services Organization staff using a central dashboard.

**Figure 1.** Outline of the intervention workflow. PHSO: Population Health Services Organization.




## Sample

Our study sample was drawn from a subset of PCP referrals that occurred between October 2020 and July 2022 (n=2512). Patients were excluded if they were nonambulatory, under 18 years of age, pregnant, institutionalized, or dependent on supplemental oxygen. The analysis focused on a subsample of referred patients with MCC. Referral timing—the intervention point—varied across the study window. To assess the program's longer-term impact rather than short-term reductions in SBP that might result from program initiation, we included SBP observations from months 7 to 12 after referral. This yielded 2206 patients with both prereferral and postreferral data for empirical analysis. Baseline data included prior BP readings, sex, age, and address from the EHR, along with self-reported race and ethnicity. Patient addresses were linked to the Healthy Places Index [34], which aggregates neighborhood-level indicators across California—such as education, employment, income, and housing—to approximate local socioeconomic conditions.

## Data Analysis

We used a semiparametric event study to assess the program's impact on SBP among patients with hypertension, focusing on those with co-occurring type 2 diabetes or ischemic heart disease, as well as those with all 3 conditions simultaneously. Our event study implementation assesses the effects of the "event" of being referred to the digital health program based on deviations of patient outcomes from prereferral trends. This design makes causal inference more complex and limited compared to a randomized clinical trial but allows us to analyze a real-world care program within a large health care system. We provide estimates for all referred patients in the event study analysis in the 2 steps described subsequently.

In the first step, we adjusted the raw outcome variables based on preintervention data using a linear time trend in calendar months and age fixed effects to account for underlying trends in patients' BP that would have occurred in the absence of the intervention. We used data from 12 to 2 months before the referral event as the baseline period. Month 1 before referral

was excluded to remove possible transitory changes preceding the referral. We then computed residuals, namely, the differences between observed SBP values in our sample and the values predicted by our model as if the intervention had not occurred.

In the second step, we examined the change in the residualized outcomes between the preevent analysis period (months 12 to 1 before the referral) to the postevent analysis period (months 7 to 12 following the referral). This analysis identified changes in SBP around the time of referral, detecting deviations from the outcome predicted by our model of patients with hypertension described in step 1, thereby indicating the impact of the program on SBP. The analysis used regressions that included patient fixed effects, which accounted for time-invariant heterogeneity in outcomes across patients. Overall, we used the comparison of SBP outcomes before and after the intervention relative to the underlying trend as the estimates for the impact of referral on BP outcomes. Significance threshold was $P<.05$.

### Ethical Considerations

This quality improvement project was deemed nonhuman subject research and exempted from institutional review board review by the UC San Diego Health Aligning and Coordinating Quality Improvement, Research, and Evaluation Committee. No compensation was provided to participants in exchange for their participation. Data were stored on encrypted UC San Diego Health servers in accordance with Health Insurance Portability and Accountability Act and UC San Diego Health policies.

## Results

Table 1 provides average characteristics for our study's population age, sex, race or ethnicity, and Healthy Places Index of residence.

**Table .** Sample characteristics (N=2206).

| | Hypertension only (n=1318) | Hypertension with either diabetes or ischemic heart disease (not both) (n=745) | Hypertension with both diabetes and ischemic heart disease (n=143) |
|---|---|---|---|
| Age (y), mean (SD) | 63.2 (15.2) | 67.1 (12.8) | 69.6 (11.5) |
| Proportion of patients aged ≥65 | 0.54 | 0.65 | 0.69 |
| HPI[a] percentile, mean (SD) | 64 (25) | 60 (26) | 59 (27) |
| Female, n (%) | 716 (54.32) | 422 (56.64) | 73 (51.05) |
| Race or ethnicity, n (%) | | | |
| Non-Hispanic White | 805 (61.08) | 343 (46.04) | 66 (46.15) |
| Black or Hispanic | 267 (20.26) | 225 (30.20) | 39 (27.27) |

[a]HPI: Healthy Places Index.

Table 2 has 3 parts. Part A presents the estimates from the full referred population. Observations from before the referral come from months −12 to −1, and observations from after the referral come from months 7 to 12. Robust SEs are clustered at the patient level. Baseline levels are reported as subpopulation means of the outcome variable in month −1 relative to referral. The results indicated notable reductions in SBP in all subpopulations. Patients with hypertension only experienced a reduction of 9.758 (SE 0.81; $P<.001$) mm Hg. For those with hypertension and either diabetes or ischemic heart disease (but not both), the reduction was 6.599 (SE 1.13; $P<.001$) mm Hg. Patients with all 3 conditions—hypertension, diabetes, and ischemic heart disease—showed a reduction of 6.604 (SE 1.73; $P<.001$) mm Hg. Baseline SBP levels at the time of referral (month t=−1) were 138.4 mm Hg for the hypertension-only group, 131.4 mm Hg for those with co-occurring hypertension and either diabetes or ischemic heart disease, and 128.82 mm Hg for those with all 3 chronic conditions.

**Table .** Program's impact on systolic blood pressure across patient groups.

| | Hypertension only | Hypertension with either diabetes or ischemic heart disease (not both) | Hypertension with both diabetes and ischemic heart disease |
|---|---|---|---|
| Part A[a,b] | | | |
| Treatment effect on referred group, mean (SE)[c] | −9.758 (0.81)[c] | −6.599 (1.13)[c] | −6.604 (1.73)[c] |
| Baseline SBP, mean (SD)[d] | 138.4 (21.4) | 131.4 (23.3) | 128.8 (22) |
| Part B[e,f] | | | |
| Initial take-up rate (SE)[c] | 0.5797 (0.0136)[c] | 0.4993 (0.0183)[c] | 0.4126 (0.0413)[c] |
| Part C | | | |
| Treatment effect on SBP scaled by take-up rate | −16.83 | −13.22 | −16.01 |

[a]Difference between hypertension only and hypertension with either diabetes or ischemic heart disease (not both) columns: 3.16 (SE 1.39; $P<.05$).

[b]Difference between hypertension with either diabetes or ischemic heart disease (not both) and hypertension with both diabetes and ischemic heart disease columns: −0.006 (SE 2.06).

[c]$P<.001$

[d]SBP: systolic blood pressure.

[e]Difference between hypertension only and hypertension with either diabetes or ischemic heart disease (not both) columns: −0.0803 (SE 0.0227; $P<.01$).

[f]Difference between hypertension with either diabetes or ischemic heart disease (not both) and hypertension with both diabetes and ischemic heart disease columns: −0.0867 (SE 0.0456; $P<.1$ ).

Part B reports the initial take-up rates of the program, which describes the number of enrollees who completed at least 1 electronic BP reading. This rate varied significantly across the patient groups. The initial take-up rate for the hypertension-only group was 57.97 percentage points (pp; SE 1.36 pp, $P<.001$). For those with hypertension and either diabetes or ischemic heart disease, the take-up rate was 49.93 pp (SE 1.83 pp, $P<.001$). Patients with all 3 conditions had a take-up rate of 41.26 pp (SE 4.13 pp, $P<.001$). The difference in initial take-up rates between those with hypertension only and those with co-occurring combinations (hypertension and either diabetes or ischemic heart disease) was −8.03 pp (SE 2.27 pp, $P<.001$), and between those with co-occurring combinations and those with all 3 conditions was −8.67 pp (SE 4.56 pp, $P=<.057$).

Part C presents the average treatment effect among participants, which was derived by dividing the treatment effect estimates from Part A by the take-up rates from Part B, thereby estimating the treatment effect for patients who enrolled in the RPM. Among program participants, the average treatment effect was a SBP reduction of 16.83 mm Hg for the hypertension-only group, 13.22 mm Hg for those with hypertension and either diabetes or ischemic heart disease, and 16.01 mm Hg for those with all 3 chronic conditions. These results indicate that the digital health program effectively reduced SBP in patients with hypertension and MCC.

# Discussion

## Principal Findings

The study results demonstrate that the team-based RPM service substantially improved BP management for patients with hypertension. This positive impact was especially notable among those with MCC, specifically patients with co-occurring diabetes or ischemic heart disease, or all 3 conditions simultaneously [7,8,35]. Although the MCC groups had a lower baseline SBP—potentially indicating more intensive prior management—this study still substantiates and extends the findings of previous prospective quality improvement programs [14] and randomized controlled trials focused on patients with hypertension alongside hyperlipidemia, diabetes [32,33], or heart disease [24]. This shows the value of RPM for managing chronic conditions in the context of a robust care team and technical support for patients who tend to be older and are more likely to have multiple conditions beyond hypertension. Indeed, there is in situ evidence that with sufficient care team and technical support, age is not a limiting factor in RPM adherence for hypertension management if patient intention to control their hypertension remains high [36]. The observed reductions in SBP were similar to those reported by comparable programs embedded within large academic health centers, which typically include ongoing care team support, multilingual care, and extensive technical resources [19]. Therefore, our findings illustrate the capacity of RPM initiatives to improve care quality.

## Implications

Concrete clinical findings such as these call for increased uptake of RPM. Most states now provide Medicaid reimbursement for RPM [21], and as the financial incentives of RPM become more apparent—1 recent RPM program for cardiovascular health saved 173% of the program cost [37]—we expect that health systems will increasingly turn to RPM programs for improved clinical outcomes. We expect the financial benefits of RPM to increase as more health centers adopt similar programs and benefit from economies of scale and accrued knowledge about best practices for establishing such programs, as well as managing patient support procedures from a technical and clinical perspective. Our EHR-integrated team-based RPM service leverages the primary care physician-patient relationship,

EHR infrastructure, and population health nurse care managers to improve clinical outcomes among patients with hypertension and MCC. It also leverages strong technical support resources from a large academic medical center. Findings from this real-world implementation of evidence bode well for further dissemination. Furthermore, reservations about patient capacity to operate technology are swiftly being addressed through the development of increasingly passive, discrete wearable sensors, such as rings [38]. Further technological developments promise an increasingly frictionless deployment of RPM programs.

### Limitations

This study has several limitations. First, while patient fixed effects were included to control for time-invariant heterogeneity, unmeasured confounders could have influenced the results. Second, for ethical reasons, the study eschewed an experimental design using randomization, as this would prevent patients randomized into a control group from receiving optimal care as recommended by their PCP. Although this represents a methodological limitation, it mirrors real-world conditions in which embedding RPM into routine care is an organizational priority, given the existing evidence supporting team-based practice assisted by RPM [39]. However, it is possible that patients referred to the program during the study period might differ systematically from those who were not in some unanticipated way. Third, the generalizability of findings from active participants is restricted to that group, as it is possible that they are more motivated, more comfortable with using digital technology, or healthier. However, findings based on all referred patients may be more broadly generalizable to referred individuals. Fourth, patient adherence varied across patient groups, which affects the ongoing effectiveness of the RPM program in addressing health disparities and equity. It is plausible that sicker patients may be less likely to adhere due to the complex tasks of managing their health. Finally, the digital health program was implemented as a team-based care model supported by technology. Importantly, this study does not attribute the observed reduction in BP solely to the technological component of RPM.

### Conclusions

This study contributes to the growing literature emphasizing that EHR-integrated RPM with a team-based management approach is associated with better clinical outcomes among patients with uncontrolled hypertension and MCC. As health systems adopt more value-based care models, programs such as these can leverage the broad adoption of EHRs, nurse care managers working alongside PCPs, and the proliferation of digital RPM devices to drive better population-level outcomes. The lower adherence rates among patients with MCC highlight a clear mandate for RPM technology development; solutions must offer maximal ease of use to effectively accommodate the complexity of managing MCC.

## Acknowledgments

The authors express gratitude for the help and guidance of clinicians and staff at University of California San Diego Health Population Health Services Organization and the Digital Health Team. No generative AI was used in this study.

## Funding

This work was supported by the University of California San Diego Health Population Health Services Organization.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

## Authors' Contributions

Conceptualization: IF, CAL, PA, MT-S
Formal analysis: IF, MT-S
Funding acquisition: MT-S
Methodology: IF, MT-S
Supervision: MT-S
Writing—original draft: RG
Writing—review and editing: CAL, PA, MT-S

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Population Health Services Organization digital health protocol for outreach and ongoing care.
[DOCX File, 16 KB - cardio_v9i1e75170_app1.docx ]

## References

1. Hypertension. World Health Organization. 2025 Sep 25. URL: https://www.who.int/news-room/fact-sheets/detail/hypertension [accessed 2025-11-24]
2. Fryar CD, Kit B, Carroll MD, Afful J. Hypertension prevalence, awareness, treatment, and control among adults age 18 and older: United States, August 2021-August 2023. NCHS Data Brief 2024 Oct(511):CS354233. [Medline: 40085792]
3. Boersma P, Black LI, Ward BW. Prevalence of multiple chronic conditions among US adults, 2018. Prev Chronic Dis 2020 Sep 17;17:E106. [doi: 10.5888/pcd17.200130] [Medline: 32945769]
4. Alecxih L, Shen S, Chan I, Taylor D, Drabek J. Individuals living in the community with chronic conditions and functional limitations: a closer look. : Office of the Assistant Secretary for Planning and Evaluation; 2010 URL: https://aspe.hhs.gov/reports/individuals-living-community-chronic-conditions-functional-limitations-closer-look-0 [accessed 2025-11-24]
5. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA 2007 Mar 21;297(11):1233-1240. [doi: 10.1001/jama.297.11.1233] [Medline: 17374817]
6. Tinetti ME, McAvay GJ, Chang SS, et al. Contribution of multiple chronic conditions to universal health outcomes. J Am Geriatr Soc 2011 Sep;59(9):1686-1691. [doi: 10.1111/j.1532-5415.2011.03573.x] [Medline: 21883118]
7. Lochner KA, Cox CS. Prevalence of multiple chronic conditions among Medicare beneficiaries, United States, 2010. Prev Chronic Dis 2013 Apr 25;10:E61. [doi: 10.5888/pcd10.120137] [Medline: 23618541]
8. Ward BW, Schiller JS. Prevalence of multiple chronic conditions among US adults: estimates from the National Health Interview Survey, 2010. Prev Chronic Dis 2013 Apr 25;10:E65. [doi: 10.5888/pcd10.120203] [Medline: 23618545]
9. Wilson PW. An epidemiologic perspective of systemic hypertension, ischemic heart disease, and heart failure. Am J Cardiol 1997 Nov 13;80(9B):3J-8J. [doi: 10.1016/s0002-9149(97)00831-x] [Medline: 9375942]
10. Agbor-Etang BB, Setaro JF. Management of hypertension in patients with ischemic heart disease. Curr Cardiol Rep 2015 Dec;17(12):119. [doi: 10.1007/s11886-015-0662-0] [Medline: 26482762]
11. Tang M, Nakamoto CH, Stern AD, et al. Effects of remote patient monitoring use on care outcomes among Medicare patients with hypertension: an observational study. Ann Intern Med 2023 Nov;176(11):1465-1475. [doi: 10.7326/M23-1182] [Medline: 37931262]
12. Margolis KL, Asche SE, Dehmer SP, et al. Long-term outcomes of the effects of home blood pressure telemonitoring and pharmacist management on blood pressure among adults with uncontrolled hypertension: follow-up of a cluster randomized clinical trial. JAMA Netw Open 2018 Sep 7;1(5):e181617. [doi: 10.1001/jamanetworkopen.2018.1617] [Medline: 30646139]
13. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote patient monitoring via non-invasive digital technologies: a systematic review. Telemed J E Health 2017 Jan;23(1):3-17. [doi: 10.1089/tmj.2016.0051] [Medline: 27116181]
14. Blood AJ, Cannon CP, Gordon WJ, et al. Results of a remotely delivered hypertension and lipid program in more than 10 000 patients across a diverse health care network. JAMA Cardiol 2023 Jan 1;8(1):12-21. [doi: 10.1001/jamacardio.2022.4018] [Medline: 36350612]
15. Walker RC, Tong A, Howard K, Palmer SC. Patient expectations and experiences of remote monitoring for chronic diseases: systematic review and thematic synthesis of qualitative studies. Int J Med Inform 2019 Apr;124:78-85. [doi: 10.1016/j.ijmedinf.2019.01.013] [Medline: 30784430]
16. Zhang Y, Peña MT, Fletcher LM, Lal L, Swint JM, Reneker JC. Economic evaluation and costs of remote patient monitoring for cardiovascular disease in the United States: a systematic review. Int J Technol Assess Health Care 2023 Apr 28;39(1):e25. [doi: 10.1017/S0266462323000156] [Medline: 37114456]
17. Hailu R, Sousa J, Tang M, Mehrotra A, Uscher-Pines L. Challenges and facilitators in implementing remote patient monitoring programs in primary care. J Gen Intern Med 2024 Oct;39(13):2471-2477. [doi: 10.1007/s11606-023-08557-x] [Medline: 38653884]
18. Lee G, Chang A, Pal A, Tran TA, Cui X, Quach T. Understanding and addressing the digital health literacy needs of low-income limited English proficient Asian American patients. Health Equity 2022 Jul 4;6(1):494-499. [doi: 10.1089/heq.2022.0045] [Medline: 36186613]
19. Schoenthaler A, Hack R, Mandal S, et al. Closing hypertension equity gaps through digitally inclusive remote patient monitoring. NEJM Catal Innov Care Deliv 2025 Feb 19;6(3). [doi: 10.1056/CAT.24.0194]
20. Bosworth HB, Olsen MK, Grubber JM, et al. Two self-management interventions to improve hypertension control: a randomized trial. Ann Intern Med 2009 Nov 17;151(10):687-695. [doi: 10.7326/0003-4819-151-10-200911170-00148] [Medline: 19920269]
21. Hayes CJ, Dawson L, McCoy H, et al. Utilization of remote patient monitoring within the United States health care system: a scoping review. Telemed J E Health 2023 Mar;29(3):384-394. [doi: 10.1089/tmj.2022.0111] [Medline: 35819861]
22. Burckhardt MA, Roberts A, Smith GJ, Abraham MB, Davis EA, Jones TW. The use of continuous glucose monitoring with remote monitoring improves psychosocial measures in parents of children with type 1 diabetes: a randomized crossover trial. Diabetes Care 2018 Dec;41(12):2641-2643. [doi: 10.2337/dc18-0938] [Medline: 30377184]
23. Jackson GL, Oddone EZ, Olsen MK, et al. Racial differences in the effect of a telephone-delivered hypertension disease management program. J Gen Intern Med 2012 Dec;27(12):1682-1689. [doi: 10.1007/s11606-012-2138-x] [Medline: 22865016]

24. Fisher ND, Fera LE, Dunning JR, et al. Development of an entirely remote, non-physician led hypertension management program. Clin Cardiol 2019 Feb;42(2):285-291. [doi: 10.1002/clc.23141] [Medline: 30582181]
25. Clark D III, Woods J, Zhang Y, Chandra S, Summers RL, Jones DW. Home blood pressure telemonitoring with remote hypertension management in a rural and low-income population. Hypertension 2021 Dec;78(6):1927-1929. [doi: 10.1161/HYPERTENSIONAHA.121.18153] [Medline: 34757773]
26. Wu JR, Cummings DM, Li Q, et al. The effect of a practice‐based multicomponent intervention that includes health coaching on medication adherence and blood pressure control in rural primary care. J Clin Hypertens (Greenwich) 2018 Apr;20(4):757-764. [doi: 10.1111/jch.13265] [Medline: 29577574]
27. Walker A, Tai-Seale M, Fadlon I, Agnihotri P, Longhurst C. Results of a team-based remote monitoring program serving patients with hypertension. . 2023 Nov p. 5179. [doi: 10.1370/afm.22.s1.5179]
28. Park SH, Shin JH, Park J, Choi WS. An updated meta-analysis of remote blood pressure monitoring in urban-dwelling patients with hypertension. Int J Environ Res Public Health 2021 Oct 9;18(20):10583. [doi: 10.3390/ijerph182010583] [Medline: 34682329]
29. Mooney MV, Guipoco FM, Gutierrez-Alvarez L, et al. Enhancing patient engagement and hypertension control rates through tailored technical assistance in community health centers. J Health Care Poor Underserved 2025;36(1):401-409. [doi: 10.1353/hpu.2025.a951606] [Medline: 39957659]
30. Patel E, Bisson C, Suresh S, et al. Systematic treatment and management of postpartum hypertension using remote patient monitoring. Pregnancy Hypertens 2025 Mar;39:101180. [doi: 10.1016/j.preghy.2024.101180] [Medline: 39709789]
31. Kuan PX, Chan WK, Fern Ying DK, et al. Efficacy of telemedicine for the management of cardiovascular disease: a systematic review and meta-analysis. Lancet Digit Health 2022 Sep;4(9):e676-e691. [doi: 10.1016/S2589-7500(22)00124-8] [Medline: 36028290]
32. Logan AG, Irvine MJ, McIsaac WJ, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57. [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
33. Logan AG, McIsaac WJ, Tisler A, et al. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. Am J Hypertens 2007 Sep;20(9):942-948. [doi: 10.1016/j.amjhyper.2007.03.020] [Medline: 17765133]
34. Maizlish N, Delaney T, Dowling H, et al. California Healthy Places Index: frames matter. Public Health Rep 2019;134(4):354-362. [doi: 10.1177/0033354919849882] [Medline: 31095451]
35. Lin PJ, Kent DM, Winn A, Cohen JT, Neumann PJ. Multiple chronic conditions in type 2 diabetes mellitus: prevalence and consequences. Am J Manag Care 2015 Jan 1;21(1):e23-e34. [Medline: 25880265]
36. Heredia NI, Fernandez ME, Garza ER, et al. Federally qualified health center patients' experiences with remote patient monitoring as part of telehealth services for self-measured blood pressure monitoring. BMC Health Serv Res 2025 Jan 20;25(1):104. [doi: 10.1186/s12913-025-12253-3] [Medline: 39833791]
37. Abraham C, Jensen C, Rossiter L, Dittman Hale D. Telenursing and remote patient monitoring in cardiovascular health. Telemed J E Health 2024 Mar;30(3):771-779. [doi: 10.1089/tmj.2023.0187] [Medline: 37682280]
38. Nwibor C, Haxha S, Ali MM, et al. Remote health monitoring system for the estimation of blood pressure, heart rate, and blood oxygen saturation level. IEEE Sensors J 2023;23(5):5401-5411. [doi: 10.1109/JSEN.2023.3235977]
39. Tai-Seale M, Downing NL, Jones VG, et al. Technology-enabled consumer engagement: promising practices at four health care delivery organizations. Health Aff (Millwood) 2019 Mar;38(3):383-390. [doi: 10.1377/hlthaff.2018.05027] [Medline: 30830826]


## Abbreviations

**BP:** blood pressure
**EHR:** electronic health record
**MCC:** multiple chronic condition
**PCP:** primary care physician
**RPM:** remote patient monitoring
**SBP:** systolic blood pressure
**pp:** percentage points

*Edited by A Coristine; submitted 29.03.25; peer-reviewed by A Tsehay, S Burchim; revised version received 12.11.25; accepted 13.11.25; published 08.12.25.*

*Please cite as:*
*Graham R, Fadlon I, Agnihotri P, Longhurst CA, Tai-Seale M*
*Outcomes of Team-Based Digital Monitoring of Patients With Multiple Chronic Conditions: Semiparametric Event Study*
*JMIR Cardio 2025;9:e75170*
*URL: https://cardio.jmir.org/2025/1/e75170*
*doi:10.2196/75170*

# Daily Dietary Sodium Intake Among Clinical Trial Participants Recruited From a University Health System or a Federally Qualified Health Center: Secondary Analysis of Baseline Participant Characteristics

Gabriella V Rubick[1], MD; Michael P Dorsch[2], PharmD, MS; Scott L Hummel[3,4], MD, MS; Tanima Basu[3], MA, MS; Evan Luff[3], MS; Kimberly Warden[5], LPN; Michael Giacalone[5], MD; Sarah Bailey[6], PhD; Mark W Newman[7], PhD; Lesli E Skolarus[8], MD, MS; Brahmajee K Nallamothu[3,9], MD, MPH; Jessica R Golbus[3,9], MD, MS

[1]Department of Internal Medicine, University of Michigan–Ann Arbor, 1500 E Medical Center Dr., Ann Arbor, MI, United States

[2]Department of Clinical Pharmacy, College of Pharmacy, University of Michigan–Ann Arbor, Ann Arbor, MI, United States

[3]Division of Cardiovascular Medicine, Department of Internal Medicine, University of Michigan–Ann Arbor, Ann Arbor, MI, United States

[4]Department of Internal Medicine, Section of Cardiology, VA Ann Arbor Healthcare System, Ann Arbor, MI, United States

[5]Hamilton Community Health Network, Flint, MI, United States

[6]Bridges into the Future, Flint, MI, United States

[7]School of Information and Department of Electrical Engineering and Computer Science, University of Michigan–Ann Arbor, Ann Arbor, MI, United States

[8]Department of Neurology, Stroke and Vascular Neurology, Northwestern Medicine, Chicago, IL, United States

[9]Michigan Integrated Center for Health Analytics and Medical Prediction, University of Michigan–Ann Arbor, Ann Arbor, MI, United States

**Corresponding Author:**
Gabriella V Rubick, MD
Department of Internal Medicine, University of Michigan–Ann Arbor, 1500 E Medical Center Dr., Ann Arbor, MI, United States

## Abstract

**Background:** Efforts to improve diversity in clinical trials often prioritize recruitment based on broad demographic factors. This approach may overlook the influence of community context and health-related social needs on health behaviors, including sodium intake, a key modifiable risk factor for hypertension and cardiovascular disease.

**Objective:** This study aims to assess the impact of enrollment site, sociodemographic factors, and health-related social needs on baseline dietary sodium intake among participants in a mobile health clinical trial aimed at lowering blood pressure.

**Methods:** The myBPmyLife study is a prospective, randomized controlled trial evaluating a mobile health intervention to lower blood pressure through increased physical activity and lower sodium food choices. Participants with hypertension were recruited from a university health system and a federally qualified health center (FQHC). All participants completed a validated sodium screener at enrollment. Sociodemographic data and health-related social needs were self-reported. Univariable and multivariable linear regression models were used to evaluate the associations between sodium intake and participant characteristics. This analysis presents a cross-sectional examination of the baseline characteristics of participants enrolled in the myBPmyLife study.

**Results:** Among 600 included participants, 96 (16.0%) were from the FQHC. Mean age was 60.1 (SD 13.5) years; 48.2% (289/600) were women, and 13.0% (78/600) were Black. FQHC participants were significantly younger (mean age 47.9, SD 11.1 vs 62.5, SD 12.7 years), more likely to be Black (43/96, 44.8% vs 35/504, 6.9%), and 8.5 times more likely to have difficulty paying for their health-related social needs. Mean baseline sodium intake was 3082.3 (SD 1072.5) mg/day, with 85.5% (513/600) of participants exceeding the World Health Organization's recommended daily sodium limit. Baseline sodium intake was significantly higher for FQHC participants (mean difference 381.1, SD 1064.2 mg/d; 95% CI 84.5-677.7; *P*=.01), men (mean difference 543.9, SD 1038.3 mg/d; 95% CI 377.3-710.5; *P*<.001), Black participants (mean difference 442.5, SD 1043.4 mg/d; 95% CI 119.7-765.3; *P*=.008) and those with difficulty affording basic needs (mean difference 338.1, SD 1066.7 mg/d; 95% CI 95.2-581.0; *P*=.02). Sodium intake was lower in older participants (−196.4 mg/d per 10 years; 95% CI −258.0 to −134.9; *P*<.001). In a multivariable analysis, age, gender, and race remained independently associated with sodium intake, while differences by site and health-related social needs were not statistically significant.

**Conclusions:** Differences in sodium intake were observed across sociodemographic groups. While the enrollment site was not independently associated with sodium intake after adjustment, it played a role in shaping the participant population, evidenced by the differences in demographics and health-related social needs among participants based on enrollment site. These findings

underscore the importance of recruiting from distinct clinical settings to capture a range of contextual factors that influence health behaviors. Clinical trials aiming for representativeness should consider both individual- and community-level factors during recruitment to more accurately inform interventions and health outcomes.

**Trial Registration:** ClinicalTrials.gov NCT05154929; https://www.clinicaltrials.gov/study/NCT05154929

***(JMIR Cardio 2025;9:e71343)*** doi:10.2196/71343

**KEYWORDS**

sodium consumption; clinical trial; hypertension; cardiovascular diseases; blood pressure; dietary sodium; sociodemographic factors

## Introduction

There is an increasing focus on recruiting diverse participants for clinical trials to ensure representative study populations [1-3]. These recruitment strategies mainly aim to include diverse populations based on standard demographic characteristics such as gender, race, age, and ethnicity. For instance, these factors have been linked to dietary sodium intake [4-11], a major contributor to hypertension and cardiovascular disease development [12,13]. However, concentrating solely on these broad participant characteristics of underrepresented groups may overlook the influence of community-level differences on health outcomes, as women, Black individuals, or older patients from one community may differ from those of similar backgrounds in another.

Although there is an increased understanding of how individual demographic factors influence health behaviors, a knowledge gap remains regarding how the context of recruitment settings and related community-level factors affect both participant traits and baseline health behaviors such as sodium intake in clinical trials. Thus, exploring differences between site characteristics and health outcomes can improve our understanding of population diversity and aid in developing targeted interventions.

This analysis contributes to this topic by analyzing a cohort of patients with hypertension recruited into the myBPmyLife clinical trial. This trial is well suited to investigate this question, as it included community members from 2 distinct clinical settings: a large, academic university health system and a federally qualified health center (FQHC) clinic. Notable population-level differences were observed between the 2 clinical settings during the trial, including the requirement for study team resources during recruitment and health-related social needs [14]. By analyzing sodium intake at baseline in a diverse participant group, we aimed to understand the impact of demographic characteristics, health-related social needs, and enrollment site on dietary sodium consumption within a mobile health (mHealth) clinical trial.

## Methods

### Study Design

The myBPmyLife study is a prospective, randomized controlled, remotely administered trial (ClinicalTrials.gov NCT05154929). Participants with hypertension were recruited from Michigan Medicine in Ann Arbor, MI, and the Hamilton Community Health Network in Flint, MI. Participants were randomly assigned to the intervention group, which received an mHealth intervention promoting increased physical activity and lower sodium food choices, or to the control group. This analysis reports on a secondary study objective: understanding sodium intake in hypertensive individuals with diverse sociodemographic characteristics. It is a baseline analysis of data collected from participants prior to randomization, and as such, all participants, regardless of study group, are included. The full study protocol and results for the primary trial outcomes have been published [15,16]. The authors are solely responsible for the design and conduct of the study, all study analyses, and drafting and editing of the paper.

### Eligibility

The study was designed to recruit patients with self-reported hypertension who could safely be physically active and reduce their sodium intake. Patients were considered eligible if they were 18 years of age or older with self-reported hypertension, had no hypertensive medication changes in the prior 4 weeks, and had a compatible smartphone. Exclusion criteria included contraindications to physical activity or sodium restriction, a secondary cause of hypertension, and a sodium consumption of <1500 mg/day as estimated by the NutritionQuest Sodium Screener (NutritionQuest), which was completed by all potential participants following informed consent. Full inclusion and exclusion criteria are available in Table S1 in Multimedia Appendix 1. In addition, 2 participants listed their gender as “other.” Due to the extremely limited sample size within this category and the resultant inability to perform meaningful statistical analyses or draw robust conclusions regarding this subgroup, these participants were excluded from this cross-sectional analysis.

### Trial Procedures

The myBPmyLife study launched in December 2021. Participants were recruited from Michigan Medicine and the Hamilton Community Health Network. Michigan Medicine is a large quaternary referral center. Its facilities are primarily located within Ann Arbor, with a median household income of US $84,245 and a 13.8% poverty rate. In contrast, the Hamilton Community Health Network is an FQHC network in Flint, MI, and delivers primary care services. Flint, MI, has a median household income of US $35,451, with a poverty rate of 33.3%, over double the poverty rate of Ann Arbor [17].

Recruitment varied by study site. The *International Classification of Diseases, Tenth Revision* (*ICD - 10*) code I10 identified participants at both sites. Potentially eligible participants were recruited using weekly emails and text

messages. Participants were preferentially recruited if they had an upcoming appointment at the Hamilton Community Health Network or with a Michigan Medicine primary care physician. The study intervention focused on delivering push notifications tailored to the participants to promote physical activity and the selection of lower-sodium food choices. The study mobile app provided feedback on participants' progress toward these goals.

### Data Collection and Study End Points

All participants were required to download the MyDataHelps mobile app, a commercially available research app developed by CareEvolution. After completing the informed consent process, participants used the app to complete the NutritionQuest Sodium Screener. Participants with an estimated sodium intake of less than 1500 mg/day were excluded from the study. The NutritionQuest Sodium Screener is a 26-item screener developed from 24-hour recall data from adults in the National Health and Nutrition Examination Survey (NHANES) from 2007 to 2008. For this screener, foods contributing to 80% of sodium intake are included, with survey respondents asked to rate the frequency at which they consume each of the foods over the prior 24 hours. This sodium screener has previously been validated in 2 studies as compared to 24-hour dietary recall with good correlation noted [18,19].

Patient-reported surveys were used to obtain sociodemographic characteristics, health-related social needs, and medical comorbidities. Specifically, health-related social needs were determined by the question, "How hard is it for you to pay for the very basics like food, housing, medical care, and heating?" Audio and video calls were used to collect medication data and to confirm sociodemographic information when necessary due to technical issues.

### Statistical Analysis

Baseline sociodemographic characteristics are described as means and SDs for continuous symmetric variables and medians with IQRs for skewed continuous variables. Categorical variables are presented as counts and percentages. We performed 2-tailed Student *t* tests to compare continuous variables and the chi-square test for categorical variables. Age and sodium intake were coded as continuous variables, with all other sociodemographic categories coded as categorical variables.

We performed a series of univariable and multivariable generalized linear models to estimate associations between key sociodemographic and clinical characteristics and sodium intake. Features within multivariable models were selected a priori based on clinical expertise and the available literature and included clinical setting, age, gender, race, and health-related social needs. Subsequently, stratified models were developed for each clinical setting. The confidence level for the lower and upper confidence limits was set at 95%. Statistical analysis was completed using Statistical Analysis System version 14.2 (SAS Institute).

### Ethical Considerations

The University of Michigan Institutional Review Board (HUM00205845) approved the study. Informed consent was obtained for all participants after the nature and possible consequences of the study were explained, in compliance with local, institutional, and national regulations on research involving human participants. Informed consent was obtained by telephone, with the option to use videoconferencing as needed, with the consent form signed within the MyDataHelps mobile app. Study data were aggregated and published in a deidentified fashion. Incentives for participation included a financial incentive of US $100 split over 2-month time points and the ability to keep the study-provided smartwatch and blood pressure cuff, contingent on survey completion and engagement with the mobile intervention for at least 2 months.

## Results

### Study Population

Between December 2021 and July 2023, 752 patients consented to participate in the myBPmyLife study. A CONSORT (Consolidated Standards of Reporting Trials) diagram for this study can be found in Figure 1. Of these 752 patients, 150 were ineligible, with 128 excluded for having a baseline sodium intake of <1500 mg/day, 13 from the FQHC, and 115 from the university health system. Participants who were ineligible due to a sodium intake of <1500 mg/day were more likely to be women (89/128; 69.5%), White (98/128; 76.6%), and not Hispanic (125/128; 97.7%) (Table S2 in Multimedia Appendix 1). After exclusions, 602 participants enrolled in the study. Two participants reported their gender as "other" and were excluded from subsequent analyses due to the limited sample size. Of the remaining 600 participants, 96/600 (16.0%) were from the FQHC. Participants' mean age was 60.1 years (SD 13.5), 289/600 (48.2%) self-identified as women, and 78/600 (13.0%) self-identified as Black. With regard to comorbidities, 28/600 (4.7%) participants reported a history of chronic kidney disease, 267/600 (44.5%) reported a history of hyperlipidemia, 118/600 (19.7%) reported a history of depression, 43/600 (71.7%) reported a history of atrial fibrillation or atrial flutter, 16/600 (2.5%) reported a history of stroke, and 134/600 (22.3%) reported a history of diabetes. FQHC participants were significantly younger than university health system participants (47.9, SD 11.1 vs 62.5, SD 12.7 years) and more likely to be Black (43/96, 44.8% vs 35/504, 6.9%) and women (61/96, 63.5% vs 228/504, 45.2%) (Table 1). In addition, 56.3% (54/96) of FQHC participants reported at least some difficulty paying for "the very basics like food, housing, medical care, and heating" compared to 6.6% (33/504) of the university health system participants.

**Figure 1.** Consolidated Standards of Reporting Trials (CONSORT) diagram for the myBPmyLife trial.

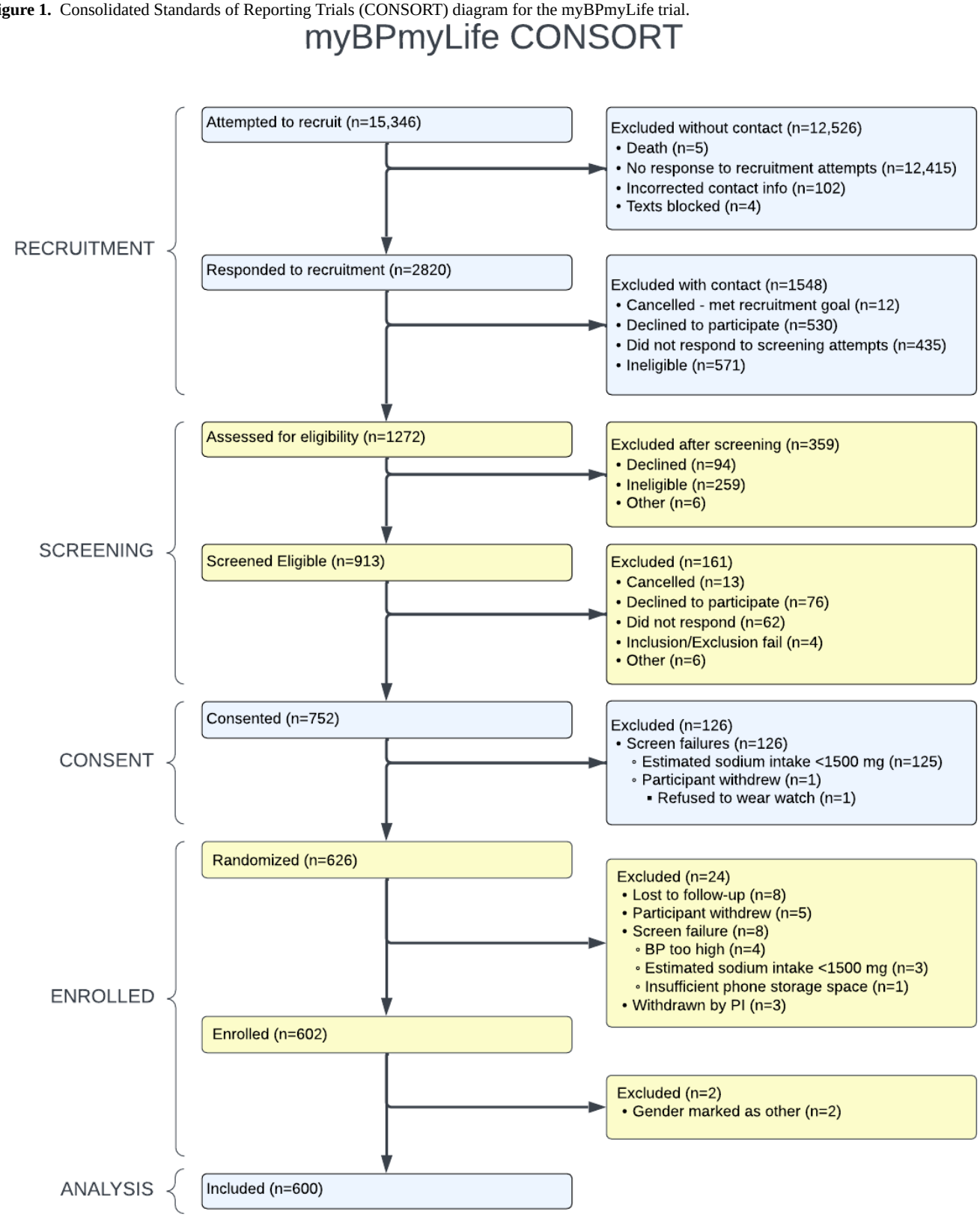


**Table .** Participant characteristics stratified by study site.

| | University health system (n=504) | FQHC[a] (n=96) | Total (N=600) | *P* value |
|---|---|---|---|---|
| Age (years), mean (SD) | 62.5 (12.7) | 47.9 (11.1) | 60.1 (13.5) | <.001 |
| Gender, n (%) | | | | .001 |
| Women | 228 (45.2) | 61 (63.5) | 289 (48.2) | |
| Men | 276 (54.8) | 35 (36.5) | 311 (51.8) | |
| Race, n (%) | | | | <.001 |
| White | 402 (79.8) | 44 (45.8) | 446 (74.3) | |
| Black | 35 (6.9) | 43 (44.8) | 78 (13.0) | |
| Other[b] | 67 (13.3) | 9 (9.4) | 76 (12.7) | |
| Ethnicity, n (%) | | | | .94 |
| Hispanic | 15 (3.0) | 3 (3.1) | 18 (3.0) | |
| Health-related social needs question, n (%) | | | | |
| How hard is it for you to pay for the very basics such as food, housing, medical care, and heating? | | | | <.001 |
| Not hard at all | 471 (93.5) | 42 (43.8) | 513 (85.5) | |
| Somewhat hard | 31 (6.2) | 47 (49) | 78 (13) | |
| Very hard | 2 (0.4) | 7 (7.3) | 9 (1.5) | |

[a]FQHC: federally qualified health center.

[b]Other: Asian, American Indian, Native Hawaiian, Other Pacific Islander, multiple, other, or refused to answer.

## Factors Associated With Elevated Sodium Intake

Most study participants consumed more than the World Health Organization–recommended 2000 mg of sodium daily (513/600, 85.5%) [20]. The study cohort's mean sodium intake was 3082.3 (SD 1072.5) mg/day, ranging from 1503 to 8886 mg/day (Table 2; Figure S1 in Multimedia Appendix 1). Sodium intake was 3021.4 (SD 987.5) mg/day for university health system participants compared to 3402.5 (SD 1402.2) mg/day for FQHC participants.

Sodium intake varied across sociodemographic characteristics. Notably, sodium intake was lower in older participants compared to younger participants. Sodium intake was significantly higher in men compared to women and in Black compared to White participants. Finally, participants who reported it was "somewhat hard" or "very hard" to pay for health-related social needs such as food, housing, medical care, and heating consumed significantly more dietary sodium than those who found it "not hard" to pay for those items(Table 2).

**Table .** Univariable analysis of sodium intake based on sociodemographic characteristics and site of enrollment.

| Characteristics | | Mean | 95% CI | *P* value |
|---|---|---|---|---|
| Clinical setting | | | | |
| | University Health System | 3021.4 (987.5) | 2934.9 to 3107.8 | 0.01 |
| | FQHC[a] | 3402.5 (1402.2) | 3118.4 to 3686.6 | N/A[b] |
| Age (per 10 years) | | −196.4 (768.9) | -258 to -134.9 | <.001 |
| Gender | | | | |
| | Women | 2800.4 (985.3) | 2686.4 to 2914.5 | <.001 |
| | Men | 3344.3 (1085.2) | 3223.2 to 3465.4 | N/A |
| Race | | | | |
| | White | 3021.8 (974.3) | 2931.1 to 3112.5 | N/A |
| | Black | 3464.3 (1376.3) | 3154 to 3774.6 | 0.008 |
| | Other[c] | 3045.6 (1197.9) | 2771.9 to 3319.4 | 0.92 |
| Ethnicity | | | | |
| | Non-Hispanic | 3082.9 (1067.6) | 3002.2 to 3176.7 | N/A |
| | Hispanic | 3064.9 (1255.6) | 2440.5 to 3689.3 | 0.92 |
| Health-Related Social Needs Question | | | | |
| | How hard is it for you to pay for the very basics like food, housing, medical care, and heating? | | | |
| | Somewhat hard/very hard | 3371.4 (1313.9) | 3091.4 to 3651.5 | 0.02 |
| | Not hard | 3033.3 (1019.3) | 2944.9 to 3121.7 | N/A |

[a]FQHC: federally qualified health center.

[b]Not applicable.

[c]Asian, American Indian, Native Hawaiian, Other Pacific Islander, multiple, other, or refused to answer.

A multivariable model was also constructed to understand the relative contributions of sociodemographic characteristics to dietary sodium intake (Table 3). As in the univariable analyses, sodium intake was lower in older individuals and was higher in men and Black participants. However, there was no significant difference in sodium intake between clinical settings or health-related social needs after accounting for these other factors (Table 3).

**Table .** Multivariate analysis of sodium intake for the entire cohort and distributed by clinical setting.

| | Entire cohort | | | | University health system | | | | FQHC[a] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Estimate | LCL[b] | UCL[c] | *P* value | Estimate | LCL | UCL | *P* value | Estimate | LCL | UCL | *P* value |
| Age (per 10 years) | −153.5 | −221 | −86.9 | <.001 | −104.4 | −172.4 | −36.5 | .003 | −450.8 | −686.8 | −214.8 | <.001 |
| How hard is it for you to pay for the very basics such as food, housing, medical care, and heating? | | | | | | | | | | | | |
| Not hard | −172.2 | −441.9 | 97.4 | .21 | −296.6 | −641.2 | 48.1 | .092 | −181.9 | −705.3 | 341.4 | .50 |
| Somewhat hard or very hard | —[d] | — | — | — | — | — | — | — | — | — | — | — |
| Race | | | | | | | | | | | | |
| Black | 311.8 | 44.4 | 579.2 | .02 | 268.9 | −61.8 | 599.6 | .11 | 413.3 | −121.2 | 947.9 | .13 |
| Other[e] | −134.1 | −382.2 | 113.9 | .29 | −77.0 | −325.3 | 171.2 | .54 | −244.8 | −1213.1 | 723.6 | .62 |
| White | — | — | — | — | — | — | — | — | — | — | — | — |
| Ethnicity | | | | | | | | | | | | |
| Non-Hispanic | 164.5 | −313.3 | 642.3 | .50 | 47.0 | −442.2 | 536.1 | .85 | 516.7 | −1037.4 | 2070.7 | .52 |
| Hispanic | — | — | — | — | — | — | — | — | — | — | — | — |
| Gender | | | | | | | | | | | | |
| Women | −565.9 | −731.0 | −400.9 | <.001 | −529.3 | −698.1 | −360.6 | <.001 | −783.8 | −1315.5 | −252.0 | .005 |
| Men | — | — | — | — | — | — | — | — | — | — | — | — |
| Clinical setting | | | | | | | | | | | | |
| FQHC | 51.9 | −231.5 | 335.2 | .72 | — | — | — | — | — | — | — | — |
| University health system | — | — | — | — | — | — | — | — | — | — | — | — |

[a]FQHC: federally qualified health center.

[b]LCL: lower confidence limit.

[c]UCL: upper confidence limit.

[d]Not applicable.

[e]Asian, American Indian, Native Hawaiian, Other Pacific Islander, multiple, other, or refused to answer.

## Discussion

### Principal Findings

Understanding the impact of demographic factors, health-related social needs, and enrollment sites on outcomes in mHealth clinical trials is relevant to ensuring enrollment of diverse patient populations. Ultimately, ensuring the recruitment of representative cohorts is crucial to shaping accurate health data for future clinical care. In this contemporary trial, sodium intake was significantly higher in younger participants, men, Black participants, and those who had difficulty paying for health-related social needs. Many of these differences in sodium intake are both statistically and clinically significant, as data show that reducing daily sodium intake by just 400 mg significantly impacts population-level cardiovascular events and mortality [21]. In addition, we found significant differences in sodium intake between our 2 study sites. However, the impact of the enrollment site on sodium intake was minimized after accounting for demographic characteristics and health-related social needs, indicating that these factors play a crucial role in study outcomes. Ultimately, this study highlights the unique influence of recruitment settings on clinical trial participant diversity and baseline health behaviors, specifically sodium intake. While the enrollment site was not independently associated with sodium intake after adjusting for individual factors, it serves as a crucial contextual indicator for underlying demographic and socioeconomic influences on health behaviors. This underscores the importance of considering community-specific factors and health-related social needs beyond broad demographic categories when recruiting for trials.

## Comparison to Prior Work

In the multivariable models of this study, only age and sex remained significantly associated with sodium intake. These findings are consistent with prior studies, including data from NHANES for the years 1999 - 2016, which have shown that younger adults and men tend to consume more sodium, in part due to differing dietary preferences and behavioral patterns [4-8,22]. While race and health-related social needs were both associated with sodium intake in unadjusted analyses, these associations attenuated in adjusted models, suggesting that their effects may be mediated or confounded by other covariates. It is also possible that our binary measure of health-related social needs did not fully capture the complexity of socioeconomic disadvantage. Notably, 1999 - 2016 NHANES data also do not show a difference in sodium intake by income level, but do note higher sodium intake in White individuals compared to Black individuals, which contrasts with our findings. Despite the lack of independent statistical significance in our adjusted models, it is possible that structural and contextual differences, such as those captured by study site, still play a meaningful role in shaping dietary behavior and access to healthy food, particularly given the large disparities observed in unadjusted comparisons.

Research into how community context affects health outcomes, known as "neighborhood effects on health," is a growing area of investigation [23]. It examines how factors such as poverty, walkability, and food accessibility affect health outcomes in a specific community and, from that, what location-specific interventions are optimal for promoting population health. Many studies show that neighborhood or community context adjustments often diminish racial and ethnic differences in health research [24-27]. Notably, in our study, 56.3% (54/96) of the FQHC cohort reported at least some difficulty paying for necessities compared to 6.6% (33/504) of university health system participants. In Flint, MI, where the FQHC is located, a large part of the community is a food desert, defined as a low-income area with at least 500 people, or 33% of the population, living more than 1 mile from the nearest supermarket [28]. In light of this, it is not surprising that individuals from the FQHC, whose primary mission is to provide care for all individuals regardless of their ability to pay for services [29], were more likely to experience difficulty in paying for health-related social needs compared to individuals from the university health system located in a comparatively more affluent area. These community factors and decreased access to healthy food options also likely affected sodium intake in FQHC participants.

## Study Strengths and Limitations

This study has multiple strengths. First, the enrollment of participants from a large university health system and an FQHC significantly augmented the diversity of our patient population. These varied enrollment sites allowed for a nuanced analysis of the role of clinical and community-based factors in health outcomes. Second, the study design, which did not necessitate in-person visits, expanded the participant pool beyond those traditionally enrolled in clinical trials. Finally, there were high response rates as baseline surveys were completed with the study coordinators.

There are several relevant limitations to this study as well. First, participants were required to own a smartphone capable of downloading study software, excluding 22 potential participants. While smartphone ownership in the United States is high (approximately 90%) [30], this criterion may still bias the study population toward more technologically resourced individuals. Future trials may consider expanding study participation through loaner phones. Second, participants were required to consume >1500 mg/day of sodium to be eligible for the study, excluding individuals who are already meeting the suggested dietary sodium goals. Third, key data, including sodium intake and sociodemographic characteristics, were self-reported. This introduces the possibility of recall bias and systematic bias, both intake-related and person-specific, as previously noted in food frequency questionnaires [31]. To mitigate this, we used a validated sodium screener. However, this screener also requires subjective reporting of dietary intake, and previous studies have noted that women tend to underreport their intake. This may account for the gender-specific differences observed in sodium intake [32]. Objective measures such as 24-hour urinary sodium or medical record review could further enhance data accuracy in future studies. Fourth, we did not adjust for total energy intake or BMI, which may partially explain higher sodium intake in certain subgroups due to greater caloric needs. However, the persistence of significant differences by race and health-related social needs suggests that other social and contextual factors may contribute to sodium intake. Fifth, we did not collect data on factors such as income, employment, insurance, or marital status, which could provide further insight into the relationship between socioeconomic factors and sodium intake. Finally, our study used a cross-sectional design, which precludes causal inference and the establishment of temporal relationships. Our findings demonstrate associations, but not causality; longitudinal studies are needed to explore causal pathways.

## Future Directions

Clinical trials aiming for representative populations must consider how site-level recruitment strategies shape participant characteristics and influence health behaviors. This study emphasizes the community-specific factors that significantly impact health-related outcomes and calls for mHealth researchers to also consider the community and clinical settings from which participants are recruited when considering diversity. In addition, future studies should collect comprehensive data on food environment and accessibility to further elucidate their impact on dietary behaviors and enhance our understanding of how these contextual factors influence trial participant characteristics and health outcomes. By doing so, we can move away from simplistic race, gender, and age-based comparisons for intervention and recruitment strategies and develop a more nuanced methodology that reflects the intricate interplay between community factors, health-related social needs, and health outcomes. This approach aligns with recent calls from the National Academies of Sciences to critically re-evaluate the use of race and ethnicity in biomedical research, advocating for a focus on underlying social and environmental determinants rather than using these categories as proxies for biological differences [33].

## Conclusions

In this clinical trial of an mHealth intervention, we observed significant differences in sodium intake among participants with variability across clinical sites and according to demographic characteristics and health-related social needs. The differences between the 2 clinical settings from which participants for this trial were recruited are limited examples, though they highlight striking differences between the groups. Although adjustment for demographic and health-related social needs minimized differences in sodium intake between sites, these broad stratifications do not fully account for the differences between these 2 communities. These findings highlight that the composition of trial populations, and thus trial outcomes, can be shaped by the recruitment setting.

## Acknowledgments

We acknowledge the wonderful staff at Hamilton Community Health Network for their help in recruitment.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

GVR wrote the original draft of this article. MPD, SH, MWN, LES, BKN, and JRG were essential to conceptualization, formal analysis, and project supervision. MPD, SH, MWN, LES, and BKN participated in funding acquisition. TB and EL conducted formal analysis. KW, MG, and SB were key project administrators and instrumental in recruiting federally qualified health center participants. All authors participated in writing, review, and editing.

## Conflicts of Interest

The study received a grant from the American Heart Association, Inc (AWD014891).
SH receives grant funding from the VA (I01CXX001636). MWN receives funding from the National Heart, Lung, and Blood Institute and the National Institute on Drug Abuse. LES receives funding from the National Institute of Aging, the National Institute of Minority Health and Health Disparities, the National Institute of Neurological Disorders and Stroke, and the American Health Association. BKN receives compensation as editor-in-chief of Circulation: Cardiovascular Quality & Outcomes, a journal of the American Heart Association. He is a co-inventor on US Utility Patent Number US15/356,012 (US20170148158A1) entitled “Automated Analysis of Vasculature in Coronary Angiograms,” which uses software technology with signal processing and machine learning to automate the reading of coronary angiograms, held by the University of Michigan. The patent is licensed to AngioInsight, Inc, in which BKN holds ownership shares and receives consultancy fees. JRG receives funding from the National Heart, Lung, and Blood Institute (L30HL143700, 1K23HL168220) and Patient-Centered Outcomes Research Institute.

Multimedia Appendix 1
Supplementary tables and figure, including full inclusion and exclusion criteria for the myBPmyLife trial, demographic characteristics of participants who were excluded due to a daily sodium intake of <1500 mg, and overall estimated dietary sodium intake (mg/day) distribution.
[DOCX File, 90 KB - cardio_v9i1e71343_app1.docx ]

## References

1. Boden-Albala B, Waddy SP, Appleton N, Kuczynski H, Nangle E, Recruitment PNS. The science of health disparities research. In: Recruitment, Inclusion, and Diversity in Clinical Trials: John Wiley & Sons; 2021:413-428. [doi: 10.1002/9781119374855]
2. Watson NL, Mull KE, Heffner JL, McClure JB, Bricker JB. Participant recruitment and retention in remote eHealth intervention trials: methods and lessons learned from a large randomized controlled trial of two web-based smoking interventions. J Med Internet Res 2018 Aug 24;20(8):e10351. [doi: 10.2196/10351] [Medline: 30143479]
3. Clark LT, Watkins L, Piña IL, et al. Increasing diversity in clinical trials: overcoming critical barriers. Curr Probl Cardiol 2019 May;44(5):148-172. [doi: 10.1016/j.cpcardiol.2018.11.002] [Medline: 30545650]
4. Abdul Aziz NS, Ambak R, Othman F, et al. Risk factors related with high sodium intake among Malaysian adults: findings from the Malaysian Community Salt Survey (MyCoSS) 2017-2018. J Health Popul Nutr 2021 May 31;40(Suppl 1):14. [doi: 10.1186/s41043-021-00233-2] [Medline: 34059146]
5. Othman F, Ambak R, Siew Man C, et al. Factors associated with high sodium intake assessed from 24-hour urinary excretion and the potential effect of energy intake. J Nutr Metab 2019;2019:6781597. [doi: 10.1155/2019/6781597] [Medline: 31192010]

6. Satheannoppakao W, Kasemsup R, Inthawong R, et al. Sodium intake and socio-demographic determinants of the non-compliance with daily sodium intake recommendations: Thai NHES IV. J Med Assoc Thai 2013 Dec;96 Suppl 5:S161-S170. [Medline: 24851587]
7. Mill JG, Malta DC, Nilson EAF, et al. Fatores associados ao consumo de sal na população adulta brasileira: pesquisa Nacional de Saúde [Article in Portuguese]. Ciênc saúde coletiva 2021;26(2):555-567. [doi: 10.1590/1413-81232021262.37492020]
8. Burnier M, Paccaud FM, Bochud M. Clinical profiles and factors associated with a low sodium intake in the population: an analysis of the Swiss Survey on Salt. Nutrients 2020 Nov 23;12(11):3591. [doi: 10.3390/nu12113591] [Medline: 33238516]
9. Hong JW, Noh JH, Kim DJ. Factors associated with high sodium intake based on estimated 24-hour urinary sodium excretion: the 2009-2011 Korea National Health and Nutrition Examination Survey. Medicine (Baltimore) 2016 Mar;95(9):e2864. [doi: 10.1097/MD.0000000000002864] [Medline: 26945369]
10. Quader ZS, Zhao L, Gillespie C, et al. Sodium intake among persons aged ≥2 years - United States, 2013-2014. MMWR Morb Mortal Wkly Rep 2017 Mar 31;66(12):324-328. [doi: 10.15585/mmwr.mm6612a3] [Medline: 28358799]
11. Fulgoni VL, Agarwal S, Spence L, Samuel P. Sodium intake in US ethnic subgroups and potential impact of a new sodium reduction technology: NHANES Dietary Modeling. Nutr J 2014 Dec 18;13(1):120. [doi: 10.1186/1475-2891-13-120] [Medline: 25522786]
12. Carriquiry A, Moshfegh AJ, Steinfeldt LC. Trends in the prevalence of excess dietary sodium intake - United States, 2003-2010. MMWR Morb Mortal Wkly Rep 2013 Dec 20;62(50):1021-1025. [Medline: 24352065]
13. Ostchega Y, Fryar CD, Nwankwo T, Nguyen DT. Hypertension prevalence among adults aged 18 and over: United States, 2017-2018. NCHS Data Brief 2020 Apr;364(364):1-8. [Medline: 32487290]
14. Skolarus LE, Farhan Z, Mishra SR, et al. Resource requirements for participant enrollment from a university health system and a Federally Qualified Health Center Network in a mobile health study: the myBPmyLife trial. J Am Heart Assoc 2024 Jan 16;13(2):e030825. [doi: 10.1161/JAHA.123.030825] [Medline: 38226521]
15. Golbus JR, Jeganathan VSE, Stevens R, et al. A physical activity and diet just-in-time adaptive intervention to reduce blood pressure: the myBPmyLife study rationale and design. J Am Heart Assoc 2024 Jan 16;13(2):e031234. [doi: 10.1161/JAHA.123.031234] [Medline: 38226507]
16. Dorsch MP, Golbus JR, Stevens R, et al. Physical activity and diet just-in-time adaptive intervention to reduce blood pressure: a randomized controlled trial. npj Digit Med 2025;8(1):438. [doi: 10.1038/s41746-025-01844-3]
17. Flint, MI & Ann Arbor, MI. Data USA. URL: https://datausa.io/profile/geo/flint-mi/?compare=ann-arbor-mi-31000US11460 [accessed 2024-07-15]
18. Norris J, Block G, Block T. Development of a brief screening tool for sodium intake. Presented at: 140th APHA Annual Meeting and Exposition; Oct 27-31, 2012; San Francisco, CA, USA.
19. Tangney CC, Rasmussen HE, Richards C, Li M, Appelhans BM. Evaluation of a brief sodium screener in two samples. Nutrients 2019 Jan 14;11(1):166. [doi: 10.3390/nu11010166] [Medline: 30646541]
20. Clarke LS, Overwyk K, Bates M, Park S, Gillespie C, Cogswell ME. Temporal trends in dietary sodium intake among adults aged ≥19 years - United States, 2003-2016. MMWR Morb Mortal Wkly Rep 2021 Oct 22;70(42):1478-1482. [doi: 10.15585/mmwr.mm7042a4] [Medline: 34673747]
21. Bibbins-Domingo K, Chertow GM, Coxson PG, et al. Reductions in cardiovascular disease projected from modest reductions in dietary salt. N Engl J Med 2010;362(7):590-599. [doi: 10.1056/NEJMoa0907355]
22. Brouillard AM, Kraja AT, Rich MW. Trends in dietary sodium intake in the United States and the impact of USDA Guidelines: NHANES 1999-2016. Am J Med 2019 Oct;132(10):1199-1206. [doi: 10.1016/j.amjmed.2019.04.040] [Medline: 31103645]
23. Arcaya MC, Tucker-Seeley RD, Kim R, Schnake-Mahl A, So M, Subramanian SV. Research on neighborhood effects on health in the United States: a systematic review of study characteristics. Soc Sci Med 2016 Nov;168:16-29. [doi: 10.1016/j.socscimed.2016.08.047] [Medline: 27637089]
24. Frieden TR. A framework for public health action: the health impact pyramid. Am J Public Health 2010 Apr;100(4):590-595. [doi: 10.2105/AJPH.2009.185652] [Medline: 20167880]
25. Diez Roux AV, Mair C. Neighborhoods and health. Ann N Y Acad Sci 2010 Feb;1186(1):125-145. [doi: 10.1111/j.1749-6632.2009.05333.x] [Medline: 20201871]
26. Morenoff JD, House JS, Hansen BB, Williams DR, Kaplan GA, Hunte HE. Understanding social disparities in hypertension prevalence, awareness, treatment, and control: the role of neighborhood context. Soc Sci Med 2007 Nov;65(9):1853-1866. [doi: 10.1016/j.socscimed.2007.05.038] [Medline: 17640788]
27. Mujahid MS, Moore LV, Petito LC, Kershaw KN, Watson K, Diez Roux AV. Neighborhoods and racial/ethnic differences in ideal cardiovascular health (the Multi-Ethnic Study of Atherosclerosis). Health Place 2017 Mar;44:61-69. [doi: 10.1016/j.healthplace.2017.01.005] [Medline: 28167269]
28. Laing S, Kato C, Kutch L, Fluegeman S. Michigan Health Endowment Fund food environment scan: a project completed by MPHI for the nutrition & healthy lifestyles program. Michigan Health Endowment Fund. 2020. URL: https://mihealthfund.org/wp-content/uploads/2020/08/HealthFund_MIFoodEnvironmentScan_202008.pdf [accessed 2025-09-15]

29. Healthcare services. Hamilton Community Health Network. URL: https://hamiltonchn.org/services/ [accessed 2024-01-19]
30. Mobile fact sheet. Pew Research Center. 2024 Jan 31. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2024-07-11]
31. Kipnis V, Subar AF, Midthune D, et al. Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol 2003 Jul 1;158(1):14-21. [doi: 10.1093/aje/kwg091] [Medline: 12835281]
32. Scagliusi FB, Ferriolli E, Pfrimer K, et al. Characteristics of women who frequently under report their energy intake: a doubly labelled water study. Eur J Clin Nutr 2009(10):1192-1199. [doi: 10.1038/ejcn.2009.5430]
33. National Academies of Sciences, Engineering, and Medicine. New report calls for changes to how biomedical research uses race and ethnicity. : The National Academies Press; 2024 URL: https://www.nationalacademies.org/news/2024/10/new-report-calls-for-changes-to-how-biomedical-research-uses-race-and-ethnicity [accessed 2025-09-22]

## Abbreviations

**CONSORT:** Consolidated Standards of Reporting Trials
**FQHC:** federally qualified health center
***ICD-10*:** *International Classification of Diseases, Tenth Revision*
**mHealth:** mobile health
**NHANES:** National Health and Nutrition Examination Survey

*Edited by A Coristine; submitted 15.01.25; peer-reviewed by D Janicki-Deverts, HC Yeh, M Kavanagh; revised version received 14.08.25; accepted 31.08.25; published 25.09.25.*

*Please cite as:*
*Rubick GV, Dorsch MP, Hummel SL, Basu T, Luff E, Warden K, Giacalone M, Bailey S, Newman MW, Skolarus LE, Nallamothu BK, Golbus JR*
*Daily Dietary Sodium Intake Among Clinical Trial Participants Recruited From a University Health System or a Federally Qualified Health Center: Secondary Analysis of Baseline Participant Characteristics*
*JMIR Cardio 2025;9:e71343*
*URL: https://cardio.jmir.org/2025/1/e71343*
*doi:10.2196/71343*

# The rs243865 Polymorphism in Matrix Metalloproteinase-2 and its Association With Target Organ Damage in Patients With Resistant Hypertension: Cross-Sectional Study

An Tuan Huynh[1], MD; Hoang Anh Vu[2], MD; Ho Quoc Chuong[2], MS; Tien Hoang Anh[3], MD; An Viet Tran[1], MD

[1]Can Tho University of Medicine and Pharmacy, 179 Nguyen Van Cu, An Khanh, Ninh Kieu, Can Tho, Vietnam

[2]Center for Molecular Biomedicine, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam

[3]Hu University of Medicine and Pharmacy, Hu , Vietnam

**Corresponding Author:**
An Viet Tran, MD
Can Tho University of Medicine and Pharmacy, 179 Nguyen Van Cu, An Khanh, Ninh Kieu, Can Tho, Vietnam

## Abstract

**Background:** Resistant hypertension (RH) presents significant clinical challenges, often precipitating a spectrum of cardiovascular complications. Particular attention recently has focused on the role of matrix metalloproteinase-2 (MMP-2) gene polymorphisms, implicated in hypertensive target organ damage (TOD). Despite growing interest, the specific contribution of MMP-2 polymorphisms to such damage in RH remains inadequately defined.

**Objective:** This study is the first to examine the rs243865 (−1306C>T) polymorphism in the MMP-2 gene in the Vietnamese population and patients with RH, underscoring its critical role as a genetic determinant of TOD.

**Methods:** A cross-sectional study with both descriptive and analytical components was conducted with 78 patients with RH at the Can Tho Central General Hospital and Can Tho University of Medicine and Pharmacy Hospital from July 2023 to February 2024.

**Results:** More than three-quarters of patients with RH had carotid-femoral pulse wave velocity (PWV) >10 m/s and microalbuminuria at a prevalence of 79% (62/78) and 76% (59/78), respectively, and more than half of patients with RH had left ventricular mass index, relative wall thickness, and carotid artery stenosis with a prevalence of 56% (45/78), 55% (43/78), and 53% (41/78), respectively. Of the 78 studied patients with RH, the presence of genotype CC was 77% (60/78), genotype CT accounted for 21% (16/78), and genotype TT for 3% (2/78). The presence of single nucleotide polymorphism rs243865 (−1306C>T) with allele T was 23% (18/78). The MMP-2 gene polymorphism 1306C/T (rs243865) was significantly associated with ejection fraction and carotid artery stenosis with odds ratios (ORs) 8.1 (95% CI 1.3 - 51.4; *P*=.03) and 4.5 (95% CI 1.1 - 20.1; *P*=.048), respectively. The allele T was found to be significantly associated with arterial stiffness including brachial-ankle PWV and carotid-femoral PWV with the correlation coefficient of OR 2.2 (95% CI 0.6 - 3.8) and OR 1.8 (95% CI 0.5 - 3.2), respectively.

**Conclusions:** The MMP-2 gene polymorphism rs243865 (−1306C>T) may have an association with measurable TOD in RH.

***(JMIR Cardio 2025;9:e71016)*** doi:10.2196/71016

**KEYWORDS**

resistant hypertension; matrix metalloproteinase-2; gene polymorphism; target organ damage; arterial stiffness

## Introduction

Resistant hypertension (RH) is characterized by the inability to achieve optimal blood pressure (BP) control despite the administration of maximum tolerated doses of antihypertensive medications, including a diuretic. This condition presents a significant clinical challenge, as it is influenced by a multitude of genetic, environmental, and pathophysiological factors that contribute to persistent hypertension. RH is closely associated with severe target organ damage (TOD), which includes damage to the heart, kidneys, and vasculature, significantly increasing the risk of cardiovascular events and mortality. Despite advancements in antihypertensive therapies, approximately 70% of patients with hypertension fail to achieve recommended BP targets, underscoring the complexity of this condition [1].

Among the molecular mechanisms contributing to RH, matrix metalloproteinases (MMPs), particularly the gelatinase family (MMP-2, MMP-9), have garnered considerable attention. These enzymes play a critical role in extracellular matrix (ECM) remodeling, a process essential to the pathogenesis of several cardiovascular diseases such as coronary artery disease, arteriosclerosis, and systemic hypertension [2]. MMP-2, in particular, has been implicated in the remodeling of cardiovascular tissues, contributing to vascular stiffness and

fibrosis, both of which are key contributors to RH and TOD [3]. Recent studies have focused on the genetic variants of the MMP-2 gene, especially single nucleotide polymorphisms (SNPs), and their potential role in the development and progression of cardiovascular diseases [4-6]. These genetic polymorphisms are believed to modulate MMP-2 expression and activity, thereby influencing the extent of cardiovascular remodeling and associated TOD. Given the growing evidence linking MMP-2 activity with hypertension-related TOD, understanding the genetic underpinnings of MMP-2 in RH could offer new insights into disease mechanisms and therapeutic targets. The objectives of this study are: (1) to investigate the clinical characteristics and extent of TOD in patients with RH; and (2) to determine the polymorphisms of the MMP-2 gene and assess their association with TOD in patients with RH.

# Methods

## Study Population

This study focused on patients with hypertension admitted to Can Tho Central General Hospital and Can Tho University of Medicine and Pharmacy Hospital from July 2023 to February 2024. The study population was divided into 2 groups: patients with RH and patients with well-controlled hypertension. The diagnosis of RH followed the 2021 guidelines of the Vietnam Hypertension Society [7].

## Sample Size

### *Overview*

To achieve the objective: "Determining the polymorphism of rs243865 and its association with TOD in patients with RH," we used the formula for estimating a single proportion. The sample size was estimated using the following formula:

$$n=Z1-a/22p(1-p)d2$$

where n=required sample size; Z=*z* score corresponding to a 95% CI (*z*=1.96); d=desired margin of error (chosen as d=0.1); and p=proportion of patients carrying the minor allele T in the RH group, estimated at 25%.

Applying the values to the formula yielded a required sample size of 72 patients with RH. In practice, 78 patients were enrolled.

### *Inclusion Criteria*

Adults aged 18 years or older diagnosed with RH, defined as the failure to achieve target BP (systolic <140 mm Hg or diastolic <90 mm Hg) despite the use of optimal or best-tolerated doses of 3 or more antihypertensive medications, including a diuretic, with BP inadequately controlled as confirmed through home or ambulatory BP monitoring, and without secondary causes of hypertension or evidence of pseudoresistant hypertension.

### *Exclusion Criteria*

Patients were excluded from the study if they had any of the following conditions: acute medical emergencies, active autoimmune diseases or ongoing immunosuppressive therapies, cancer or other malignant conditions, secondary hypertension confirmed by clinical and laboratory examinations, pregnancy or chronic kidney disease (CKD), or if they refused to participate or demonstrated nonadherence to the medication regimen.

## Methodological Approach

### *Design Framework*

The study used a cross-sectional, descriptive-analytic design to investigate the association between the SNP rs243865 (−1306C>T) in the MMP-2 gene and RH versus nonresistant hypertension. Patients were recruited from 2 hospitals from July 2023 to February 2024. Patients were classified into resistant and nonresistant hypertension groups according to the European Society of Cardiology criteria for RH.

### *Sampling Strategy*

Nonprobability convenience sampling method was used. Patients meeting inclusion criteria were recruited consecutively upon admission to the cardiology and internal medicine departments. Trained research assistants approached patients daily, explained the study objectives, and obtained informed consent prior to enrollment. Convenience sampling was selected due to logistical feasibility and time constraints.

## Research Protocol and Variables

### *Demographic and Risk Factors*

Data were systematically collected regarding the following risk factors and comorbid conditions, clearly defined based on standard clinical criteria:

- Diabetes mellitus: defined as having a documented diagnosis of diabetes, or current use of antidiabetic medications, or fasting plasma glucose ≥126 mg/dL, or $HbA_{1c}$≥6.5%.
- Overweight or obesity: defined according to BMI classification, with overweight as BMI≥25 kg/m² and obesity as BMI≥30 kg/m², calculated from measured height and weight.
- Smoking status: categorized as smoker (currently smoking ≥1 cigarette per day or having ceased smoking for at least 6 mo prior to enrollment), or nonsmoker (no lifetime smoking).
- History of heavy drinking: defined according to the National Institute on Alcohol Abuse and Alcoholism guidelines as consumption of ≥14 drinks per week for men or ≥7 drinks per week for women, or a documented history of alcohol use disorder.

These data were obtained through structured patient interviews and cross-verified by medical records to ensure accuracy and consistency.

### *Clinical and Hemodynamic Parameters*

#### Overview

BP and pulse pressure were measured using the BOSO ABI-100 system in all patients to minimize errors, with measurements taken at least twice in a seated position after 5 minutes of rest; pulse pressure was calculated as the difference between systolic and diastolic BP [8]. A 24-hour ambulatory BP monitoring device was used to assess mean systolic and diastolic BP,

nocturnal dipping, and early morning BP surge. The resting heart rate was measured manually or with a digital monitor. Blood samples were collected to determine serum levels of urea, creatinine, and electrolytes, including sodium, potassium, and chloride. TOD was evaluated across several key organs, with specific diagnostic criteria used to define damage in each organ system.

#### Cardiac Damage

Left ventricular hypertrophy (LVH) was assessed using echocardiography, with the left ventricular mass index (LVMI) calculated. According to the European Society of Cardiology guidelines, LVH was defined as LVMI >95 g/m² for women and LVMI >115 g/m² for men. Electrocardiogram criteria for LVH, such as the Sokolow-Lyon and Cornell voltage criteria, were also used as secondary diagnostic tools [1].

Left ventricular ejection fraction (EF), a key indicator of cardiac function, was measured via echocardiography. EF was classified as normal (≥50%), mildly reduced (41% - 49%), moderately reduced (30% - 40%), or severely reduced (<30%). All the echocardiography is made via Siemens Acuson X300 ultrasound machine.

#### Brain Damage

Brain damage was assessed through imaging techniques, including computed tomography and magnetic resonance imaging. The presence of any of the following conditions was considered indicative of brain damage: white matter lesions, cerebral microbleeds, lacunar infarctions, and dilated perivascular spaces.

A history of stroke or transient ischemic attack was also considered as evidence of brain damage.

#### Renal Damage

Renal damage was assessed using the urinary albumin-to-creatinine ratio. This method evaluates kidney function by measuring albumin excretion in the urine.

Renal damage was defined as an albumin-to-creatinine ratio of: normal to mildly increased (<30 mg/g); moderately increased (30 - 300 mg/g); and severely increased (>300 mg/g).

Patients with a history of CKD stage 4 or 5, or renal failure (estimated glomerular filtration rate<30 mL/min/1.73 m²), were excluded from the study to avoid confounding factors related to advanced renal failure.

#### Vascular Damage

Vascular stiffness was assessed using pulse wave velocity (PWV), defined as the speed at which arterial pressure waves move along the vessel wall, with a PWV >10 m/s being indicative of vascular damage via the BOSO ABI-100 system. The ankle-brachial index (ABI) was also measured using the BOSO ABI-100 system. ABI is defined as the ratio of the systolic BP measured at the ankle to the systolic BP measured at the brachial artery. An ABI of ≤0.9 was indicative of peripheral arterial disease and thus considered a sign of vascular damage.

Carotid artery damage was assessed using ultrasound to measure carotid intima-media thickness. Carotid stenosis was defined as the presence of plaques that caused a ≥50% reduction in the arterial lumen or if the intima-media thickness was ≥0.9 mm. Significant stenosis was confirmed through Doppler ultrasound via Siemens Acuson X300 ultrasound machine.

### MMP-2 Gene Polymorphism Analysis

#### *Sequencing Technique*

A 4 mL blood sample was collected into ethylenediaminetetraacetic acid–coated tubes and stored at 2 °C until used for DNA extraction and analysis. The SNP genotype was determined using 2 direct sequencing methods.

#### *Principle*

The sequencing technique was carried out using an automated sequencer based on a modified Sanger method. In this method, the dideoxynucleotide triphosphates are not radioactively labeled but are tagged with different fluorescent dyes for each type of dideoxynucleotide triphosphate. The automated sequencer comprises key components such as a capillary system, a laser illumination system, and a signal detection and processing system. The capillary electrophoresis bands emit light as they pass through a laser beam, and the color detection system records and encodes the nucleotides as A, T, C, or G.

#### *Main Steps in Sequencing*

DNA extraction was performed using the Qiagen DNA extraction kit (Qiagen, Hilden, Germany), following the manufacturer's protocol. The target region containing the SNP was then amplified via polymerase chain reaction (PCR). The PCR products were visualized through agarose gel electrophoresis, and subsequently purified using the Qiagen PCR purification kit (Qiagen, Hilden, Germany). Sequencing of the purified PCR products was carried out using the modified Sanger method. Capillary electrophoresis was performed on a Beckman Coulter CEQ8000 sequencer. The sequence data were further analyzed using the ABI 3500 Genetic Analyzer (Applied Biosystems, Foster City, California, United States). Sequence processing and SNP analysis were conducted with SeqScape software (version 2.7; Applied Biosystems). The results were interpreted by comparing the identified SNP locations with the corresponding reference sequences retrieved from the National Center for Biotechnology Information database.

#### *Method*

Sequencing was performed using the Beckman Coulter CEQ8000 sequencer.

### Statistical Analysis

The dataset underwent statistical treatment using Stata (version 15.1; StataCorp) and was articulated through frequency distribution (for qualitative variables), and mean (SD; for quantitative measures). Comparison for qualitative data was made by chi-square tests and by 2-tailed Student *t* tests for quantitative data. A significance level of .05 was used for all tests to establish statistical significance. Stepwise multiple regression analysis with inclusion at the .01 level was used to evaluate the influence of gen rs243865 (−1306C>T) on targeted organ damage adjusted by clinical and subclinical characteristics. To estimate the relationship between MMP-2

gene SNPs and TOD, odds ratio and its 95% CI were calculated for binary TOD variables including echocardiogram EF and carotid artery stenosis. Regression coefficients (β reg coef.) and its 95% CI were calculated for continuous TOD variables including brachial-ankle PWV (m/s) and carotid-femoral PWV (cfPWV; m/s). The squared correlation coefficient ($R^2$) was calculated for the proportion of variance explained by the model.

### Ethical Considerations

The study was approved by the Ethics Council in Biomedical Research, Can Tho University of Medicine and Pharmacy, through the research ethics approval form 23.006.NCS/HĐĐĐ dated June 15, 2023, before data collection. The study was also licensed to be conducted at Can Tho Central General Hospital and Can Tho University of Medicine and Pharmacy Hospital. The study was conducted with the consent of the participants through the consent form. The process of interview and the implementation of testing techniques were conducted conveniently and comfortably for the participants, not related to private issues that may affect the health or psychology of the participants. Participants did not receive any compensation for their participation. The personal information of the participants was kept confidential. This study aimed to protect and improve public health and has no other purpose.

## Results

The protocol is presented in the study diagram (Figure 1). In our analysis of 78 patients with RH, a significant proportion were female (49/78, 63%), with an average age of 66.7 (SD 14.4) years. The majority of patients (51/78, 65%) were older than 60 years of age, highlighting the predominance of an older cohort. Notably, 68% (53/78) of the patients had a history of hypertension extending beyond 10 years, reflecting the chronic nature of RH, which complicates BP control (Table 1).

**Figure 1.** Study protocol. BP: blood pressure.

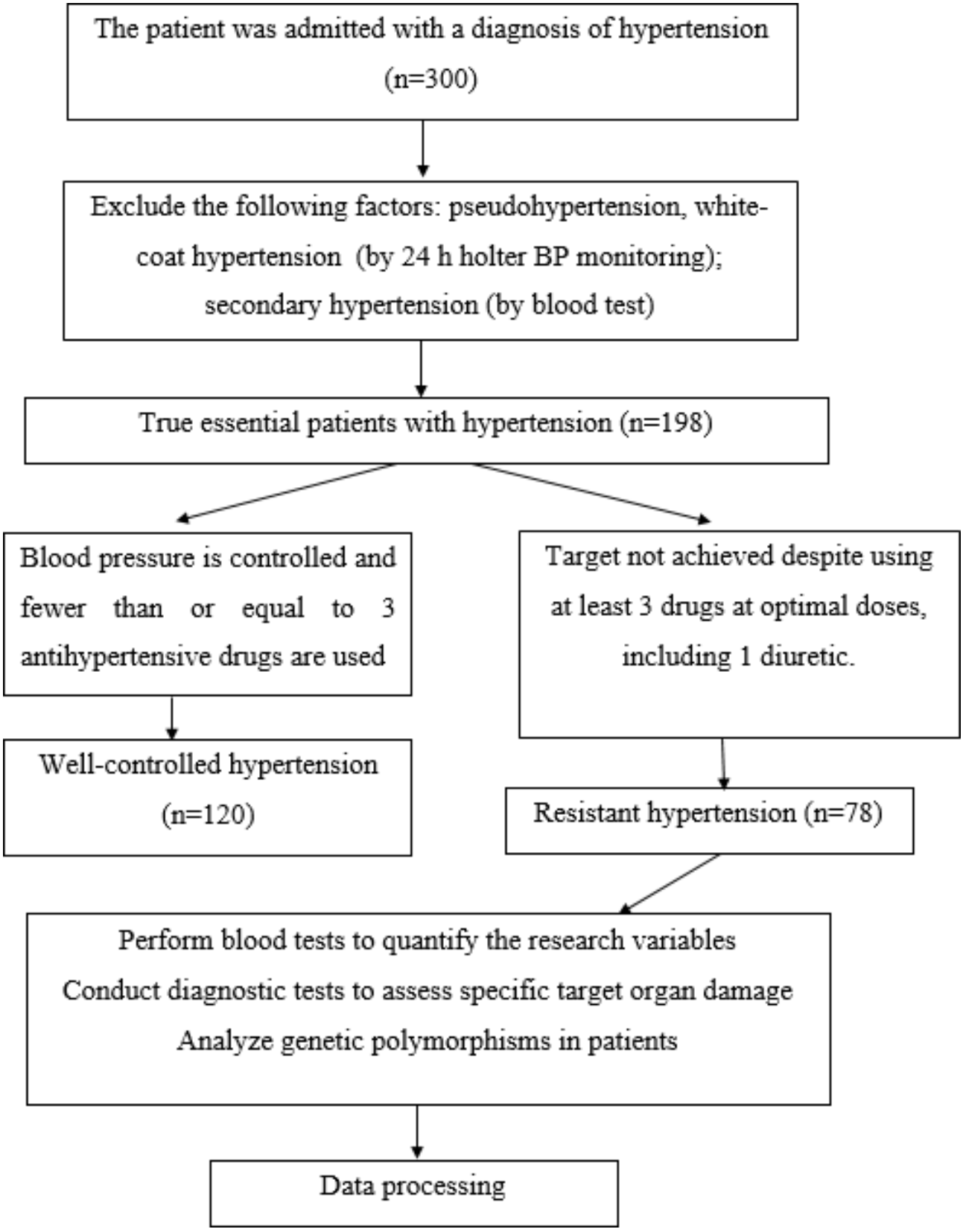


**Table .** Clinical characteristics of patients with RH[a].

| Clinical characteristics | Value (N=78), n (%) |
|---|---|
| Sex | |
| Male | 29 (37) |
| Female | 49 (63) |
| Age[b] (years) | |
| ≤60 | 27 (35) |
| >60 | 51 (65) |
| Duration of hypertension[c] (years) | |
| ≤10 | 53 (68) |
| >10 | 25 (32) |
| Blood pressure level | |
| Grades 1 and 2 | 53 (68) |
| Grade 3 | 25 (32) |
| Diabetes | |
| Yes | 22 (28) |
| No | 56 (72) |
| Overweight or obese | |
| Yes | 20 (26) |
| No | 58 (74) |
| Smoking (current or past history) | |
| Yes | 24 (31) |
| No | 54 (69) |
| History of heavy drinking | |
| Yes | 25 (32) |
| No | 53 (68) |
| Triglyceride[d] (mmol/L) | |
| ≥2.26 | 38 (49) |
| <2.26 | 40 (51) |
| LDL[e] (mmol/L)[f] | |
| ≥3.36 | 24 (31) |
| <3.36 | 54 (69) |
| Blood lipid disorders | |
| Yes | 49 (63) |
| No | 29 (37) |

[a]RH: resistant hypertension.

[b]Age: mean 66.7 (SD 14.4) years.

[c]Duration of hypertension: mean 10.3 (SD 5.6).

[d]Triglyceride: mean 2.85 (SD 2.42)

[e]LDL: low-density lipoprotein.

[f]LDL: mean 2.95 (SD 1.28)

Despite treatment adherence, mean systolic and diastolic BP levels were persistently elevated, averaging 162.5 (SD 29.6) mm Hg and 92.7 (SD 15.9) mm Hg, respectively. This underscores the therapeutic challenges posed by RH. Common comorbidities included diabetes (22/78, 28%) and obesity (20/78, 26%). Additionally, dyslipidemia was prevalent, with

high serum triglycerides (38/78, 49%) and low-density lipoprotein cholesterol (24/78, 31%). The prevalence of TOD was striking, with 79% (62/78) of patients demonstrating cfPWV >10 m/s, an indicator of increased arterial stiffness. Microalbuminuria, found in 76% (59/78) of patients, suggests significant renal impairment, while over half of the cohort showed elevated LVMI and increased relative wall thickness, both markers of adverse cardiac remodeling driven by chronic hypertension (Table S1 in Multimedia Appendix 1).

The MMP-2 gene polymorphism rs243865 (−1306C>T) was investigated, revealing that 77% (60/78) of patients carried the CC genotype, while 21% (16/78) carried the CT genotype, and 3% (2/78) the TT genotype (Table 2). The T allele frequency was 23% (18/78), potentially highlighting a genetic predisposition for more severe vascular outcomes in RH.

**Table .** Distribution of MMP-2[a] gene polymorphism rs243865 (−1306C>T) in patients with RH[b].

| MMP-2 gene polymorphism rs243865 (−1306C>T) | Value (N=78), n (%) |
|---|---|
| Genotype | |
| CC | 60 (77) |
| CT | 16 (21) |
| TT | 2 (3) |
| Allele | |
| T carrier | 18 (23) |
| CC | 60 (77) |

[a]MMP-2: matrix metalloproteinase-2.

[b]RH: resistant hypertension.

Significant relationships were identified between the T allele and specific TOD markers, particularly reduced EF and increased cfPWV. T allele carriers exhibited a lower mean EF (53.8, SD 20.3) compared to noncarriers (62.1, SD 12.7), with a statistically significant difference (*P*=.04). Additionally, T allele carriers had higher brachial-ankle PWV and cfPWV values, nearing statistical significance (both *P*=.07), suggestive of enhanced arterial stiffness (Table 3).

**Table .** The comparison mean of target organ damage indicators between MMP-2[a]–carrying polymorphisms nucleotide at rs243865 (-1306C>T) with and without allele T.

| Indicators of target organ damage | T carrier (n=18), mean (SD) | CC (n=60), mean (SD) | *P* value[b] |
|---|---|---|---|
| Left ventricular mass index (g/m$^2$) | 120.1 (55.9) | 114.9 (44.9) | .69 |
| EF[c] in echocardiogram | 53.8 (20.3) | 62.1 (12.7) | .04 |
| Blood pressure difference | 70.3 (15.5) | 71.4 (22.1) | .84 |
| ABI[d] | 0.98 (0.15) | 0.99 (0.2) | .76 |
| Brachial-ankle PWV[e] (m/s) | 19.1 (3.5) | 17.4 (3.5) | .07 |
| Carotid-femoral PWV (m/s) | 13.6 (2.9) | 12.2 (2.9) | .07 |
| eGFR[f] | 66.6 (27.2) | 74.4 (32.3) | .36 |
| ACR[g] | 130.2 (147.7) | 140.5 (182.9) | .84 |

[a]MMP-2: matrix metalloproteinase-2.

[b]*P* value: independent samples 2-tailed *t* test.

[c]EF: ejection fraction.

[d]ABI: ankle-brachial index.

[e]PWV: pulse wave velocity.

[f]eGFR: estimated glomerular filtration rate.

[g]ACR: albumin-to-creatinine ratio.

The association between the T allele and carotid artery stenosis was also notable, with 72% (13/18) of T allele carriers exhibiting stenosis compared to 47% (28/60) of noncarriers, approaching statistical significance (*P*=.06; Table 4). T allele carriers exhibited a higher prevalence of EF of <40% and carotid artery stenosis compared to noncarriers (Table 5). Specifically, 22% (4/18) of T allele carriers had an EF of <40%, compared to only 7% (4/60) of noncarriers, approaching statistical significance (*P*=.06). Similarly, carotid artery stenosis was present in 72% (13/18) of T allele carriers versus 47% (28/60) of noncarriers

(*P*=.06), indicating a potential role of the T allele in exacerbating arterial remodeling and stenosis (Table 4). After adjusting for age and serum potassium levels, the T allele remained significantly associated with EF <40% (Table 5). After adjusting for age, hypertension duration, and sodium levels, T allele carriers had a significantly higher risk of carotid artery stenosis (Table S2 in Multimedia Appendix 1).

**Table .** The comparison of the percentage of hypertension-mediate organ damage between MMP-2[a] polymorphisms nucleotide at rs243865 (−1306C>T) with and without allele T.

| Symptoms of target organ damage | T carrier (n=18), n (%) | CC (n=60), n (%) | *P* value[b] |
|---|---|---|---|
| History of stroke or TIA[c] | 4 (22) | 14 (23) | .92 |
| ECG[d] ischemia | 9 (50) | 18 (30) | .12 |
| ECG left ventricular hypertrophy | 4 (22) | 13 (22) | .96 |
| Echocardiogram EF[e] <40% | 4 (22) | 4 (7) | .06 |
| Echocardiogram with regional hypokinesis | 6 (33) | 22 (38) | .79 |
| Echocardiographic left ventricular mass index (>95 for women and >115 for men) | 9 (50) | 35 (58) | .53 |
| Echocardiographic relative wall thickness ≥0.43 | 10 (56) | 33 (55) | .97 |
| Carotid artery stenosis | 13 (72) | 28 (47) | .06 |
| Ankle-brachial index <0.9 | 3 (17) | 11 (18) | .87 |
| Carotid-femoral pulse wave velocity >10 m/s | 16 (89) | 45 (75) | .21 |
| eGFR[f] <60 mL/min/1.73m$^2$ | 7 (39) | 17 (28) | .39 |
| Albuminuria (urine albumin/creatinine ratio >30 µg/g) | 14 (78) | 45 (75) | .81 |

[a]MMP-2: matrix metalloproteinase-2.

[b]*P* value: chi-square.

[c]TIA: transient ischemic attack.

[d]ECG: electrocardiogram.

[e]EF: ejection fraction.

[f]eGFR: estimated glomerular filtration rate.

**Table .** Association of MMP-2[a] gene polymorphism rs243865 (−1306C>T) and echocardiogram EF[b] in resistant hypertension (N=78).

| | EF <40% | EF ≥40% | Univariate logistic regression | | Multivariate logistic regression[c] | |
|---|---|---|---|---|---|---|
| | | | OR[d] (95% CI) | *P* value | OR (95% CI) | *P* value |
| rs243865 (−1306C>T), n (%) | | | | .06 | | .03 |
| T Carrier | 4 (22) | 14 (78) | 4.0 (0.9-18.0) | | 8.1 (1.3 - 51.4) | |
| CC | 4 (7) | 56 (93) | —[e] | | — | |
| Age group (years), n (%) | | | | .09 | | .06 |
| ≤60 | 5 (19) | 22 (82) | — | | — | — |
| ≥61 | 3 (6) | 48 (94) | 0.3 (0.06 - 1.3) | | 0.2 (0.03 - 1.1) | |
| Potassium serum concentration, mean (SD) | 3.3 (0.3) | 3.6 (0.4) | 0.13 (0.14 - 1.2) | .06 | 0.1 (0.01 - 1.3) | .07 |

[a]MMP-2: matrix metalloproteinase-2.

[b]EF: ejection fraction.

[c]The 3-factor model $R^2$=0.2306.

[d]OR: odds ratio.

[e]Not applicable.

The T allele was also associated with higher cfPWV, a marker of arterial stiffness and a predictor of cardiovascular events (Table S3 in Multimedia Appendix 1). The multivariate regression model showed a significant correlation between the T allele and increased PWV (β=1.8, 95% CI 0.5 - 3.2; *P*=.008). This highlights the potential role of the rs243865 polymorphism in promoting arterial stiffness.

## Discussion

### Principal Findings

In this study, we selected patients with true RH, excluding those with advanced-stage CKD and secondary hypertension. This ensured that the TOD observed was specific to patients with primary hypertension, a population that typically receives inadequate screening for TOD. Our patient cohort, representing the health care setting of a resource-limited country, included a predominantly lower-income population. These patients often exhibit limited concern for their health and lack access to regular check-ups compared to those in high-income countries. Our findings, which were largely anticipated, emphasize several critical characteristics and clinical implications of RH. These include the difficulty in controlling BP, its association with comorbidities, and the significant burden of TOD, consistent with prior studies over the past 5 years.

### Comparison to Prior Work

#### *Demographic and Clinical Characteristics*

The predominance of female patients (49/78, 63%) and older patients (51/78, 65% older than 60 years of age) is consistent with previous research showing that RH is more prevalent among older adults and female patients [9,10]. A history of hypertension exceeding 10 years in 68% (53/78) of patients reflects the chronic nature of the condition, which not only complicates BP management but also elevates the risk of TOD [11].

Despite adherence to treatment, mean systolic and diastolic BP levels remained high (162.5, SD 29.6 mm Hg and 92.7, SD 15.9 mm Hg, respectively). This highlights the challenges of achieving BP targets in RH, which may be attributed to inflammatory mechanisms and hyperactivity of the sympathetic nervous system and the renin-angiotensin-aldosterone system [1].

The high prevalence of diabetes (22/78, 28%) and obesity (20/78, 26%) in this cohort aligns with well-established risk factors for RH. These conditions not only contribute to endothelial dysfunction but also exacerbate arterial stiffness, worsening hypertension [12,13]. Dyslipidemia, characterized by elevated triglycerides (38/78, 49%) and low-density lipoprotein cholesterol (24/78, 31%), further increases cardiovascular risk and TOD [14]. Although diabetes and obesity are not considered primary causes of secondary hypertension, effective management of weight and glucose levels can improve BP control and overall prognosis in patients with RH.

#### *TOD*

The burden of TOD in patients with RH was substantial. A high proportion of patients 79% (62/78) demonstrated elevated cfPWV (>10 m/s), indicating significant arterial stiffness—a critical marker of vascular aging and cardiovascular risk [15]. While cfPWV is predominantly used in research settings rather than routine clinical practice, it remains a robust prognostic indicator independent of brachial BP. Interestingly, we observed that cfPWV does not always correlate with BP levels, suggesting that relying solely on BP measurements may overlook high-risk patients with significant arterial stiffness. The high prevalence of elevated cfPWV in this study could be both a consequence of prolonged hypertension and a contributing factor to RH.

Microalbuminuria was observed in 76% (59/78) of patients, indicating early renal dysfunction and its central role in RH pathophysiology via sodium retention and renin-angiotensin-aldosterone system activation [11,16]. While most clinicians rely on creatinine levels and estimated glomerular filtration rate to assess renal damage, our findings reveal a concerning rate of early kidney damage even in patients without advanced CKD, warranting greater clinical attention.

LVH and increased relative wall thickness were observed in over half of the patients, consistent with previous studies highlighting the importance of echocardiography in accurately assessing cardiac TOD. Compared to electrocardiograms, echocardiography has significantly higher sensitivity in detecting LVH [16-18].

Furthermore, RH has been shown to substantially increase the risk of severe cardiovascular events, including heart failure, myocardial infarction, and stroke, particularly in ambulatory RH cases [14].

#### *Association of SNP With TOD*

Our analysis demonstrates a strong association between the rs243865 (−1306C>T) polymorphism in the MMP-2 gene and TOD in patients with RH. The results emphasize that the T allele (the minor allele) significantly increases the risk of arterial stiffness, carotid artery stenosis, and reduced EF. Previous studies have shown that rs243865 enhances the transcriptional activity of MMP-2, leading to excessive ECM degradation, which contributes to vascular and cardiac fibrosis [19,20].

In this study, cfPWV, a key indicator of arterial stiffness, was on average 1.8 m/s higher in the T allele group compared to the CC genotype group. This aligns with previous finding [21], which highlighted the critical role of MMP-2 in promoting arterial fibrosis, particularly in older individuals. Other studies also indicated that MMP-2 polymorphisms are associated with increased arterial stiffness in hypertensive populations [22,23]. Furthermore, inflammation and oxidative stress interact with MMP-2 activity, exacerbating arterial stiffness in patients with RH [24]. Evidence from multiple studies indicates that arterial stiffness is independently linked to genetic factors, irrespective of BP control, paving the way for its potential as a predictive marker for resistance to antihypertensive therapy [3,21,24].

The prevalence of carotid artery stenosis was significantly higher in the T allele group, underscoring its critical role in vascular remodeling. Our findings are consistent with previous studies, which demonstrated that rs243865 upregulates MMP-2, promoting the development of atherosclerotic plaques and narrowing the arterial lumen [19,25]. Additionally, ECM

remodeling mediated by MMP-2 reduces arterial elasticity and contributes to carotid artery stenosis [26]. However, prior studies emphasized that beyond rs243865, other genetic and environmental factors play a critical role, reflecting the multifactorial nature of this pathology [27].

Patients carrying the T allele exhibited significantly lower EF, with an average reduction of approximately 8% compared to the CC genotype group, indicating impaired cardiac function and an increased risk of heart failure. Previous studies have reported that haplotypes in the MMP-2 gene are associated with LVH, myocardial infarction, and impaired cardiac function [28,29]. The enhanced activity of MMP-2 driven by rs243865 leads to ECM degradation, destabilizing cardiac structure and triggering compensatory fibrosis. This finding presents a potential therapeutic application, as the inhibition of MMP-2 has been shown to improve cardiac function in preclinical models [30]. From a broader perspective on causality, reduced EF often originates from pressure overload and vascular remodeling. The influence of the MMP-2 gene on vascular structure, leading to arterial stiffness, may impair cardiac function by increasing afterload [21].

### *The Role of Genetics in TOD*

This study, aligned with previous studies, highlights the significant role of the rs243865 (−1306C>T) polymorphism in the MMP-2 gene in the risk of TOD [31]. This genetic variant not only exerts its effects independently but also interacts intricately with other factors such as inflammation and environmental influences. Specifically, this polymorphism increases the risks of arterial stiffness, carotid artery stenosis, and impaired cardiac function in patients with RH. Genetic variants within the MMP-2 gene can significantly alter the risk of cardiovascular diseases [5,23]. These variants play a pivotal role in vascular remodeling, leading to severe outcomes such as LVH and reduced cardiac pumping capacity. The rs243865 polymorphism, through enhanced MMP-2 activity, disrupts ECM integrity, thereby contributing to the structural weakening of the vasculature and heart [32]. Furthermore, rs243865 has been implicated in other vascular diseases beyond hypertension, including ischemic stroke and aneurysms. This underscores its potential as a critical risk factor in systemic vascular conditions. The overactivation of MMP-2 associated with rs243865 leads to excessive ECM degradation, weakening vascular structures and promoting the development and progression of vascular lesions [4,33]. Recently, intermediate factors, such as obesity and insufficient physical activity, proved capable of amplifying the effects of rs243865 on BP and TOD [6]. Obesity, through mechanisms of chronic inflammation and endocrine disruption, exacerbates MMP-2 activity, while sedentary lifestyles further contribute to vascular dysfunction [27]. Synthesizing all these findings, rs243865 emerges as not only a key genetic determinant of TOD but also a nexus of complex interactions with other factors, including inflammation, oxidative stress, lifestyle, and environmental influences. This highlights its potential as a target for personalized treatment strategies aimed at regulating MMP-2 activity and mitigating its associated impacts in the management of RH.

## Limitations

This study is limited by its small sample size, cross-sectional design, and focus on a single ethnic population, which may affect the generalizability of the findings. Additionally, unmeasured confounding factors, such as inflammation and interactions with other genetic polymorphisms, were not assessed. Further longitudinal and multiethnic studies are needed to validate these results and explore the broader implications of rs243865 and TOD in RH. First, this study used a relatively small sample size (N=78), which may limit the generalizability and statistical power of our findings. To mitigate this, we calculated the sample size based on a statistically valid estimation formula to ensure adequate representation; however, larger multicenter studies would enhance statistical power. Second, the cross-sectional design of this study prevents us from establishing a causal relationship between the rs243865 polymorphism and TOD. While this design enabled the identification of associations, longitudinal studies would be necessary to clarify causality and the temporal sequence of events. Third, although this is the first study about rs243865 in Vietnamese people, the focus on a single ethnic group limits the external validity of the findings, potentially restricting applicability to other populations. To address this, future research should include diverse ethnic groups to assess whether these genetic associations hold across different populations. Finally, due to limited data availability, we were unable to compare the genotype distribution of rs243865 in our patients with RH with that in the general Vietnamese population. This limitation should be addressed in future population-based studies to provide a more comprehensive interpretation of the genetic findings.

## Future Directions

Future research could expand the scope by exploring additional genetic polymorphisms within the MMP-2 gene and their combined impact with rs243865 on RH and associated TOD. Translating findings from genetic associations into clinical practice represents a significant opportunity. Genetic screening for MMP-2 polymorphisms could facilitate personalized medicine approaches by identifying patients at higher risk for RH and severe TOD, allowing clinicians to initiate more aggressive or targeted interventions earlier in the treatment course. Additionally, therapeutic strategies targeting MMP-2 activity, such as the use of specific inhibitors, may offer new avenues for managing and mitigating vascular and cardiac complications in patients with RH and patients with cardiovascular disease as in our prior study [34].

## Conclusions

This study underscores the critical role of the rs243865 (−1306C>T) polymorphism in the MMP-2 gene as a significant genetic determinant of TOD in patients with RH. Our findings highlight the multifaceted impact of this polymorphism, including its association with increased arterial stiffness, carotid artery stenosis, and reduced EF. Importantly, the influence of rs243865 extends beyond its direct genetic effects, interacting with inflammation, oxidative stress, and modifiable factors such as obesity and physical activity. The high prevalence of TOD in our patient population underscores the urgent need for

comprehensive screening and management strategies, particularly in resource-limited settings where access to advanced diagnostic tools remains a challenge.

The study provides compelling evidence for considering rs243865 as a potential biomarker for risk stratification and a target for therapeutic intervention. Future research should focus on validating these findings in larger and more diverse populations, exploring the mechanistic pathways linking MMP-2 activity to TOD, and evaluating the clinical efficacy of MMP-2 inhibitors in reducing vascular and cardiac complications in patients with RH. Moreover, integrating genetic testing for rs243865 into clinical practice could pave the way for personalized treatment approaches, allowing for more targeted and effective management strategies.

## Acknowledgments

We thank the patients and health care professionals at Can Tho Central General Hospital and Can Tho University of Medicine and Pharmacy for their participation and support. We also appreciate the guidance from the Research Ethics Committee and the contributions of our colleagues in facilitating this study.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

## Authors' Contributions

ATH contributed to the conceptualization, formal analysis, funding acquisition, investigation, methodology, supervision, and writing—original draft. HQC contributed to the data curation, formal analysis, genetic sequencing, investigation, methodology, software, and validation. HAV contributed to the resources, validation, genetic sequencing, and supervision. THA contributed to the investigation, methodology, validation, and writing—review and editing. AVT contributed to the project administration, conceptualization, supervision, validation, visualization, and writing—review and editing.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File, 17 KB - cardio_v9i1e71016_app1.docx ]

## References

1. McEvoy JW, McCarthy CP, Bruno RM, et al. 2024 ESC guidelines for the management of elevated blood pressure and hypertension. Eur Heart J 2024 Oct 7;45(38):3912-4018. [doi: 10.1093/eurheartj/ehae178] [Medline: 39210715]
2. Henriet P, Emonard H. Matrix metalloproteinase-2: not (just) a “hero” of the past. Biochimie 2019 Nov;166:223-232. [doi: 10.1016/j.biochi.2019.07.019] [Medline: 31362036]
3. Peeters SA, Engelen L, Buijs J, et al. Circulating matrix metalloproteinases are associated with arterial stiffness in patients with type 1 diabetes: pooled analysis of three cohort studies. Cardiovasc Diabetol 2017 Oct 25;16(1):139. [doi: 10.1186/s12933-017-0620-9] [Medline: 29070037]
4. Li Y, Ouyang QR, Li J, et al. Correlation between matrix metalloproteinase-2 polymorphisms and first and recurrent atherosclerotic ischemic stroke events: a case–control study. J Internet Med Res 2021 Jun;49(6):3000605211022967. [doi: 10.1177/03000605211022967]
5. Misra S, Talwar P, Kumar A, et al. Association between matrix metalloproteinase family gene polymorphisms and risk of ischemic stroke: a systematic review and meta-analysis of 29 studies. Gene 2018 Sep 25;672:180-194. [doi: 10.1016/j.gene.2018.06.027] [Medline: 29906531]
6. Ritter AMV, de Faria AP, Barbaro NR, et al. The rs243866/243865 polymorphisms in MMP-2 gene and the relationship with BP control in obese resistant hypertensive subjects. Gene 2018 Mar 10;646:129-135. [doi: 10.1016/j.gene.2017.12.023] [Medline: 29288728]
7. Van Minh H, Van Huy T, Long DPP, Tien HA. Highlights of the 2022 Vietnamese Society of Hypertension guidelines for the diagnosis and treatment of arterial hypertension. J Clin Hypertens (Greenwich) 2022 Sep;24(9):1121-1138. [doi: 10.1111/jch.14580] [Medline: 36196473]
8. Fendrik K, Biró K, Endrei D, et al. Oscillometric measurement of the ankle-brachial index and the estimated carotid-femoral pulse wave velocity improves the sensitivity of an automated device in screening peripheral artery disease. Front Cardiovasc Med 2023;10:1275856. [doi: 10.3389/fcvm.2023.1275856] [Medline: 38155988]
9. Bansal S. Revisiting resistant hypertension in kidney disease. Curr Opin Nephrol Hypertens 2024 Sep 1;33(5):465-473. [doi: 10.1097/MNH.0000000000001002] [Medline: 38726750]

10. Chan KK, Chiang L, Choi CC, Li Y, Chen CX. Prevalence and associated risk factors of resistant hypertension among Chinese hypertensive patients in primary care setting. BMC Prim Care 2024 Apr 19;25(1):120. [doi: 10.1186/s12875-024-02366-9] [Medline: 38641566]
11. Mishchenko LA, et al. Features of target organ damage in patients with resistant arterial hypertension. Hypertension 2021;13(2-3):66-74. [doi: 10.22141/2224-1485.13.2-3.2020.205339]
12. Bakris G, Lin PP, Xu C, Chen C, Ashton V, Singhal M. Prediction of cardiovascular and renal risk among patients with apparent treatment-resistant hypertension in the United States using machine learning methods. J Clin Hypertens (Greenwich) 2024 May;26(5):500-513. [doi: 10.1111/jch.14791] [Medline: 38523465]
13. Barbaro NR, Fontana V, Modolo R, et al. Increased arterial stiffness in resistant hypertension is associated with inflammatory biomarkers. Blood Press 2015 Feb;24(1):7-13. [doi: 10.3109/08037051.2014.940710] [Medline: 25061978]
14. Coccina F, Salles GF, Banegas JR, et al. Risk of heart failure in ambulatory resistant hypertension: a meta-analysis of observational studies. Hypertens Res 2024 May;47(5):1235-1245. [doi: 10.1038/s41440-024-01632-8] [Medline: 38485774]
15. Rizzoni D, Agabiti-Rosei C, Boari GEM, Muiesan ML, De Ciuceis C. Microcirculation in hypertension: a therapeutic target to prevent cardiovascular disease? J Clin Med 2023 Jul 25;12(15):4892. [doi: 10.3390/jcm12154892] [Medline: 37568294]
16. Bult MM, van de Ree TF, Wind AM, Hurley KM, van de Ree MA. The use of echocardiography compared to electrocardiogram when screening for left ventricular hypertrophy in hypertensive patients: a cross-sectional study. J Clin Hypertens (Greenwich) 2024 Aug;26(8):977-985. [doi: 10.1111/jch.14867] [Medline: 38963706]
17. Modolo R, de Faria AP, Paganelli MO, et al. Predictors of silent myocardial ischemia in resistant hypertensive patients. Am J Hypertens 2015 Feb;28(2):200-207. [doi: 10.1093/ajh/hpu140] [Medline: 25063735]
18. García MIP, Del Pino Y Pino MD, Rodriguez RA, Rodelo-Haad C, Carreño TP. The value of ABPM and subclinical target organ damage parameters in diagnosis of resistant hypertension. Nefrologia (Engl Ed) 2019;39(1):67-72. [doi: 10.1016/j.nefro.2018.03.007] [Medline: 30001849]
19. Hernandez NAG, et al. A polymorphism in the matrix metalloproteinase-2 (MMP-2-1306T>C) gene promoter is associated with high risk of ischemic stroke in hypertensive patients. Genes and Genomics 2008;30(6):533-540 [FREE Full text]
20. Price SJ, Greaves DR, Watkins H. Identification of novel, functional genetic variants in the human matrix metalloproteinase-2 gene. J Biol Chem 2001 Mar 9;276(10):7549-7558. [doi: 10.1074/jbc.M010242200] [Medline: 11114309]
21. Diaz-Canestro C, Puspitasari YM, Liberale L, et al. MMP-2 knockdown blunts age-dependent carotid stiffness by decreasing elastin degradation and augmenting eNOS activation. Cardiovasc Res 2022 Jul 27;118(10):2385-2396. [doi: 10.1093/cvr/cvab300] [Medline: 34586381]
22. Belo VA, Luizon MR, Carneiro PC, et al. Effect of metabolic syndrome risk factors and MMP-2 genetic variations on circulating MMP-2 levels in childhood obesity. Mol Biol Rep 2013 Mar;40(3):2697-2704. [doi: 10.1007/s11033-012-2356-7] [Medline: 23242659]
23. Sabbatini AR, Barbaro NR, de Faria AP, et al. Matrix metalloproteinase-2 -735C/T polymorphism is associated with resistant hypertension in a specialized outpatient clinic in Brazil. Gene 2017 Jul 15;620:23-29. [doi: 10.1016/j.gene.2017.04.004] [Medline: 28390988]
24. Rubio TA, Rodrigues B, da Costa LMC, et al. Linear and non-linear analyses of autonomic modulation in uncontrolled and controlled elderly resistant hypertensives. Exp Gerontol 2022 Mar;159:111686. [doi: 10.1016/j.exger.2021.111686] [Medline: 34995726]
25. Liu O, Xie W, Qin Y, et al. MMP-2 gene polymorphisms are associated with type A aortic dissection and aortic diameters in patients. Medicine (Baltimore) 2016 Oct;95(42):e5175. [doi: 10.1097/MD.0000000000005175] [Medline: 27759651]
26. Vasků A, Goldbergová M, Izakovicová Hollá L, et al. A haplotype constituted of four MMP-2 promoter polymorphisms (-1575G/A, -1306C/T, -790T/G and -735C/T) is associated with coronary triple-vessel disease. Matrix Biol 2004 Jan;22(7):585-591. [doi: 10.1016/j.matbio.2003.10.004] [Medline: 14996438]
27. Zou GY, Deng YS, Lu KY, Zeng D, Liu L, Yang Y. Association analysis between genetic variants of matrix metalloproteinase enzyme 2 gene and the blood pressure of children and adolescents. Zhonghua Xin Xue Guan Bing Za Zhi 2022 Oct 24;50(10):1000-1006. [doi: 10.3760/cma.j.cn112148-20211012-00878] [Medline: 36299223]
28. Lacchini R, Jacob-Ferreira ALB, Luizon MR, et al. Common matrix metalloproteinase 2 gene haplotypes may modulate left ventricular remodelling in hypertensive patients. J Hum Hypertens 2012 Mar;26(3):171-177. [doi: 10.1038/jhh.2011.8] [Medline: 21307882]
29. Pérez-Hernández N, Vargas-Alarcón G, Martínez-Rodríguez N, et al. The matrix metalloproteinase 2-1575 gene polymorphism is associated with the risk of developing myocardial infarction in Mexican patients. J Atheroscler Thromb 2012;19(8):718-727. [doi: 10.5551/jat.11817] [Medline: 22785600]
30. Matsumura S, Iwanaga S, Mochizuki S, Okamoto H, Ogawa S, Okada Y. Targeted deletion or pharmacological inhibition of MMP-2 prevents cardiac rupture after myocardial infarction in mice. J Clin Invest 2005 Mar;115(3):599-609. [doi: 10.1172/JCI22304] [Medline: 15711638]
31. Sabbatini A, De Faria AP, Ruggeri N, et al. [PP.23.11] Matrix metalloproteinase 2 gene polymorphism associates with resistant hypertension. J Hypertens (Los Angel) 2016;34(Supplement 2):e257. [doi: 10.1097/01.hjh.0000492082.48991.8c]

32. Sattari M, Hassanzad M, Jamaldini SH, et al. Association between matrix metaloproteinases 2-1306C/T polymorphism and the risk of coronary artery disease in Iranian population. Pathophysiology 2017 Sep;24(3):185-189. [doi: 10.1016/j.pathophys.2017.05.001] [Medline: 28522356]
33. Yu W, Peng J, Chen Z, et al. Association of plasma MMP-2 levels and prognosis of patients with intracerebral hemorrhage: a prospective cohort study. Front Neurol 2023;14:1259339. [doi: 10.3389/fneur.2023.1259339] [Medline: 38090262]
34. Tran AV, Nguyen NTK, Pham NTN, Tran BLT, Huynh AT, Ngo TH. Characteristics and predictors of cardiovascular events related to CYP2C19 gene polymorphisms following acute coronary syndrome. Eur Rev Med Pharmacol Sci 2024 Oct;28(19):4347-4354. [doi: 10.26355/eurrev_202410_36829] [Medline: 39436079]

## Abbreviations

**ABI:** ankle-brachial index
**BP:** blood pressure
**cfPWV:** carotid-femoral pulse wave velocity
**CKD:** chronic kidney disease
**ECM:** extracellular matrix
**EF:** ejection fraction
**LVH:** left ventricular hypertrophy
**LVMI:** left ventricular mass index
**MMP:** matrix metalloproteinase
**PCR:** polymerase chain reaction
**PWV:** pulse wave velocity
**RH:** resistant hypertension
**SNP:** single nucleotide polymorphism
**TOD:** target organ damage

*Edited by A Coristine; submitted 08.01.25; peer-reviewed by L Saremi, M Saberi-Karimian; revised version received 01.04.25; accepted 01.04.25; published 01.05.25.*

*Please cite as:*
*Tuan Huynh A, Vu HA, Chuong HQ, Anh TH, Viet Tran A*
*The rs243865 Polymorphism in Matrix Metalloproteinase-2 and its Association With Target Organ Damage in Patients With Resistant Hypertension: Cross-Sectional Study*
*JMIR Cardio 2025;9:e71016*
*URL: https://cardio.jmir.org/2025/1/e71016*
*doi:10.2196/71016*

# English- and Spanish-Speaking Patient Preferences on Home Blood Pressure Monitors in an Urban Safety Net Setting: Qualitative Study

Jonathan J Shih[1*], BS; Vivian E Kwok[1*], MPH; Isabel Luna[2,3], BA; Hyunjin Cindy Kim[2,3], MPH; Faviola Garcia[2,3], BA; Christian Gutierrez[2,3], BS; Mahal Garcia[3], BA; Courtney R Lyles[3,4,5], PhD; Elaine C Khoong[2,3], MD, MS

[1]School of Medicine, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, CA, United States

[2]Department of Medicine, Division of General Internal Medicine, San Francisco General Hospital, University of California, San Francisco, 2540 23rd Street, Room 4708, San Francisco, CA, United States

[3]UCSF Action Research Center for Health Equity, San Francisco General Hospital, San Francisco, CA, United States

[4]Center for Healthcare Policy and Research, University of California, Davis, Davis, CA, United States

[5]Department of Public Health Science, University of California, Davis, Davis, CA, United States

[*]these authors contributed equally

**Corresponding Author:**
Elaine C Khoong, MD, MS
Department of Medicine, Division of General Internal Medicine, San Francisco General Hospital, University of California, San Francisco, 2540 23rd Street, Room 4708, San Francisco, CA, United States

## Abstract

**Background:** Self-measured blood pressure monitoring is necessary for successful management of hypertension. However, disparities in blood pressure control persist, with low-income patients and racial and ethnic minorities more likely to have uncontrolled hypertension. These patients are also at increased risk for digital exclusion. Several validated blood pressure monitors for self-measured monitoring are available, but little is known about patient preferences between different device traits. Studies have shown that poor usability or technology design can lead to barriers to adoption.

**Objective:** We investigated patient-reported barriers, preferences, and facilitators to self-measured blood pressure monitoring from a diverse population at an urban safety-net hospital.

**Methods:** This qualitative study included English- and Spanish-speaking patients with hypertension. Participants completed a survey about sociodemographic traits, self-measured blood pressure monitoring practices and training, and experience with technology. Semi-structured interviews were conducted to elicit preferences about blood pressure devices, the accompanying mobile apps, and their experience sharing blood pressure measurements with their providers. Interviews included participant demonstration of home blood pressure measurement to evaluate baseline self-measured blood pressure monitoring technique. Two home blood pressure monitoring devices were presented: a Bluetooth-enabled device and a cellular-enabled device that syncs data directly. Surveys and interviews were conducted in participants' preferred language. Rapid qualitative data analysis was applied to analyze qualitative data.

**Results:** Fifteen participants (8 English-speaking and 7 Spanish-speaking) were enrolled. Participants all identified as racial and ethnic minorities. Educational attainment varied, ranging from less than high school to college graduates. Eight exhibited some form of digital inaccessibility: lacking internet access, not activating their patient portal, or having difficulty connecting a device to Wi-Fi. Most required assistance with Bluetooth pairing and navigating app features. Overall, participants valued tracking their blood pressure, were motivated to engage in self-measured blood pressure monitoring practices, and desired training. Nearly all participants demonstrated inconsistencies in blood pressure education, displayed incorrect measurement techniques, and had not received formal training on self-measured blood pressure monitoring. Spanish-speaking participants reported that using apps was challenging because they were presented in English and wanted translated apps and resources. The cost of features was a key factor in device preference.

**Conclusions:** Patient-reported barriers to successful self-measured blood pressure monitoring adoption include cost, insufficient training, digital inaccessibility, and language discordance. Addressing these challenges may enhance the adoption of self-measured blood pressure monitoring in safety net populations. Providers should evaluate patients' preferences and develop tailored interventions when recommending self-measured blood pressure monitoring. Cellular self-measured blood pressure monitoring devices that automatically transmit blood pressure readings may reduce digital complexity and promote sharing results with providers, though future studies are needed to evaluate usability and implementation.

*(JMIR Cardio 2025;9:e60196)* doi:10.2196/60196

**KEYWORDS**

telemedicine; telehealth; monitoring; blood pressure monitoring; hypertension; patient preference; healthcare disparities; safety-net providers; physiologic; ambulatory

## Introduction

Hypertension is highly prevalent, resulting in significant cardiovascular disease morbidity and mortality [1-3]. Self-measured blood pressure monitoring (SMBP) is an evidence-based guideline-recommended strategy for improving hypertension management [4-6], in which patients regularly check their own blood pressure (BP) at home and share these results with their clinical care team to improve hypertension control. This approach has shown promising results in populations with worse hypertension outcomes, such as individuals with low income or from racial and ethnic minority backgrounds [7-11].

Despite its potential impact, SMBP faces implementation challenges, especially in populations and health care centers where hypertension disparities are most pronounced. These populations, including Black, Latine/Hispanic, and Asian adults; those with limited insurance; individuals with lower educational attainment; and those with limited English proficiency have worse BP control [8,9,12,13].

Few studies have investigated patient perspectives on challenges to and preferences for SMBP monitoring [14-16], especially in safety net health systems where providers "organize and deliver a significant level of health care and other health-related services to uninsured, Medicaid, and other vulnerable patients" [17]. In addition, these studies do not explore in-detail preferences for home BP monitors and do not assess home measurement technique. Perspectives from patients receiving care in safety net settings are crucial for designing effective SMBP interventions tailored to communities experiencing disparities in BP control. To address the gap in knowledge on patient-reported barriers to SMBP, preferences for different types of BP monitors, and facilitators that would support SMBP in a racially diverse and low-income population, we conducted a qualitative, observational study of patients with hypertension receiving care at an urban safety net hospital.

## Methods

### Study Sample

Participants included English- and Spanish-speaking patients with a diagnosis of hypertension receiving care from an urban academic safety net system. Participants were recruited from August 2022 through October 2023 from an outpatient registry of patients with hypertension. Patients were called and invited to share information about their use of home BP monitors. Participant interviews were not connected to hospital visits. Patients were excluded if primary care clinicians thought the patients would be unable to provide consent due to cognitive impairment or capacity. Participants were recruited until data saturation was reached, as no new themes emerged during the interviews [18]. This study follows the SRQR (Standards for Reporting Qualitative Research) guidelines [19].

### Survey and Interview Administration

All participants completed a survey and a semistructured interview in their preferred language with a native speaker. Researchers followed an interview guide (Multimedia Appendix 1 and Multimedia Appendix 2), which was updated based on iterative review and feedback from the study team and included participant demonstration of home BP measurement to evaluate baseline SMBP technique. The survey included closed-ended questions about sociodemographic traits, SMBP practices and training, and experience with technology. Participants self-reported demographic information, including gender, race and ethnicity, highest educational attainment, and preferred language. A survey question was added mid-study to further assess participants' digital literacy. Interviews asked open-ended questions about BP device preferences, experience with BP device mobile apps, and current practices around tracking and sharing BP measurements with their clinicians. Each participant was presented with 2 home BP monitoring devices during the interview: a Bluetooth-enabled device (Device 1) and a cellular-enabled device that syncs data directly (Device 2). Participants were allotted 10 minutes per device to test and record a BP reading. Interviews were audio-recorded, and interviewers simultaneously documented field notes to capture participants' nonverbal cues, key observations, and immediate reflections on the interview content.

### Analysis

Rapid qualitative data analysis was used to analyze qualitative data [20-23]. Compared to traditional qualitative methods, this is a rigorous, pragmatic approach to analyzing qualitative data on an accelerated timeline. The data collection and analysis goals of rapid qualitative analysis focus on identifying key descriptors to address time-sensitive questions as opposed to developing a theoretically extensive understanding of a concept. As this study sought to inform intervention elements for future home BP monitoring trials [24-26], rapid qualitative analysis with purposeful data reduction activities was used. Rapid qualitative data analysis steps include developing a codebook of domains from interviews, summarizing interviews based on domains, and validating across the study team for consistency.

Analysis involved an iterative and systematic approach to identifying key insights from the interview data. An interview guide (Multimedia Appendices 1 and 2) was initially developed to gather feedback relevant to intervention design, which was iteratively updated to ensure subsequent interviews captured all relevant perspectives and content areas. To standardize interview analyses, a summary template was structured around 8 initial coding domains that were developed according to these emerging topics: (1) Existing general knowledge of and prior training about blood pressure, (2) Impact of device's physical

features on patient preference, (3) Considerations and technique when measuring their blood pressure at home, (4) Patient preferences: Bluetooth versus cellular connection, (5) Logging and sharing BP data, (6) User app experience: Viewing BP results and navigating features, (7) Evaluation of provided training and learning materials, and (8) Ideal training and resource preferences.

Following this initial coding and summarization process, we engaged in a second phase of analysis focused on data reduction and theme generation. Through iterative team discussions, we collapsed the 8 domains into 3 higher-order codes, each representing a major conceptual area of patient experience with SMBP. Specifically, domains 1, 7, and 8 were merged into Code 1: Patient Knowledge about SMBP. Domains 3, 5, and 6 were consolidated into Code 2: Patterns and Challenges in Self-Monitoring Blood Pressure. Domains 2 and 4 formed Code 3: Varied Patient Preferences on Home Blood Pressure Monitoring Device Features. Within each analytic code, we then identified themes, defined as recurrent, explanatory patterns that captured participants' reported barriers, preferences, and facilitators for SMBP adoption. This multistep process allowed us to move systematically from coded content to broader thematic insights. After this process, we selected representative quotes from the interview transcripts to illustrate key concepts within each code.

Three investigators (JS, VK, and IL) reviewed audio recordings and field notes independently for an initial subset of interviews. Using the template with the set of 8 domains, each interview was summarized to extract key concepts. The process continued for 8 interviews until consensus was reached on the approach and content of the interview summaries. The remaining interviews were then analyzed and summarized by a single investigator using the established framework (IL). After completing summaries from all interviews, the content within each domain was reviewed across all interviews to identify overarching themes. Particular attention was given to differences in responses by language to capture any nuanced variations. Agreement on key themes was reached through multiple collaborative discussions with the entire research team until consensus was achieved. Frequency of participant responses within each theme for these nonsurvey interview data was synthesized according to the following nomenclature: "few" representing approximately 3-5 participants, "many" representing 5-7, "most" representing the majority or at least 8 of the 15 total participants, and "nearly all" representing 13 or 14.

Analytical credibility and trustworthiness were enhanced through multiple strategies [27]. We used investigator triangulation with 3 independent coders (JS, VK, and IL) analyzing initial interviews and coordinated with clinical staff members who had years of experience onboarding patients to SMBP studies. Two senior study investigators (CL and EK) have conducted SMBP research for more than 5 years, providing prolonged engagement with the research topic. Interviews lasted more than 90 minutes and incorporated multiple follow-up questions to ensure depth of data collection and participant discussion. Emerging domains were incorporated into the codebook through iterative discussion and team consensus, allowing the analysis to evolve and reflect new insights gained throughout the process.

To ensure dependability and confirmability, we maintained comprehensive audit trails throughout the research process. This included detailed field notes from each interview, multiple versions of our codebook documenting its evolution, and iterations of the interview guide as it was refined based on emerging insights. Meeting minutes from team discussions captured analytical decisions and rationale for theme development. To enhance confirmability, multiple team members independently reviewed the coding process to ensure findings were grounded in participant data rather than researcher assumptions. At least 2 researchers reviewed each transcript, with discrepancies resolved by a third investigator or the whole research team. This audit trail and multireviewer approach helped ensure consistency of findings over time and minimized individual bias in interpretation [27]. Reflexivity was maintained through field notes documenting the researcher's observations, reactions, and preliminary interpretations during each interview. Research team meetings included structured self-reflection on how our professional backgrounds as healthcare researchers and clinicians in safety-net settings might influence data interpretation. Team members explicitly discussed potential biases related to our experiences with health disparities and digital health implementation. These discussions were used to ensure interpretations remained grounded in participant data.

### Ethical Considerations

All participants provided written informed consent and received a US $50 gift card for participation. Data were deidentified for analysis and reporting. Participants were given the option to withdraw at any time throughout the study and have their data removed. This study was approved by the University of California, San Francisco Institutional Review Board (number 21 - 33711).

## Results

### Participant Demographics and Baseline Experience With Technology and Self-Measured Blood Pressure

Table 1 displays the characteristics of 15 total participants, all of whom completed both the survey and interview. Eleven participants were assessed for their SMBP technique against a rubric using the updated interview guide. The median age of participants was 57 (IQR 37-71) years.

From survey responses, all participants had a smartphone, 10 used mobile apps several times a day, and 10 reported no difficulty using their smartphone or installing apps without assistance. Eight participants either did not have internet service at home other than via smartphone, did not know how to connect a device to Wi-Fi, or had not activated their patient portal account.

**Table .** Participant demographics and experience with technology.

| Characteristics | Values, n (%; N=15) |
|---|---|
| Female | 9 (60) |
| Race and Ethnicity | |
| American Indian or Native American | 1 (7) |
| Asian or Pacific Islander | 1 (7) |
| Black or African American | 4 (27) |
| Hispanic or Latine | 8 (53) |
| Two or more | 1 (7) |
| Highest educational attainment | |
| Less than high school | 5 (33) |
| High school graduate or GED[a] | 5 (33) |
| College graduate or more | 5 (33) |
| Preferred language | |
| English | 8 (53) |
| Spanish | 7 (47) |
| Experience with technology | |
| Frequency of using apps for any purpose on phone | |
| Several times a day | 10 (67) |
| At least once a day | 3 (20) |
| Once a week | 2 (13) |
| Difficulty installing apps on phone | |
| Not difficult | 10 (67) |
| Somewhat/Very difficult | 5 (33) |
| Has internet service at home other than via smartphone | 13 (87) |
| Activated patient portal account | 8 (53) |
| Difficulty using phone without someone else's help | |
| Not difficult | 12 (80) |
| Somewhat/Very difficult | 3 (20) |
| Knows how to connect device to Wi-Fi, (N=11)[b] | 6 (55) |

[a]GED: General Educational Development.

[b]This question was later added to the survey to further assess digital literacy

In the past 12 months, 5 participants measured their BP outside of the clinic, all at home using their own BP monitor. Overall, 3 measured their BP less than once a month, 2 measured at least once a month, and 2 shared these BP measurements with their clinical team. Four participants reported that measuring BP at home and sharing results with their clinician was "extremely helpful."

Through our analysis, we consolidated the 8 initial coding domains into 3 high-order analytical codes. Within each analytical code, we identified themes that captured the key barriers, preferences, and facilitators for SMBP adoption. All analytical codes, themes, and notable quotes are provided in Table 2.

**Table .** Analytical codes, themes, and notable quotes from interviews.

| Themes | Description | Notable quotes |
|---|---|---|
| Analytical Code 1: Patient Knowledge about SMBP[a] | | |
| Theme 1.1 Inconsistencies and Gaps Exist in Patient Education | Few participants received adequate education on home BP[b] monitoring techniques | • "Drawings and pictures help people understand faster…People never really go through books [of instructions], something simpler would be better." |
| Theme 1.2: Patients Desire Training and Education | All participants desired more in-person training and education, and in particular take-home materials that were easy to understand and language-concordant | • "[There is] a lot of writing… [I would] need to read more than once." |
| Theme 1.3: Incorrect Technique and Missing Considerations in Home BP Measurement | Few participants engaged in all recommended practices for accurate home measurement | • "It's impossible to get accurate reading at home." |
| | | • "It would've taken me a bit more time [without a demonstration]." |
| Analytical Code 2: Patterns and Challenges in Self-Monitoring Blood Pressure | | |
| Theme 2.1 Participants Like the Idea of Tracking BP | While most participants did not track their BP, they thought it was important and liked the apps' ability to do so | • "Yes it is [important to keep a log of your BP readings]. You can talk to your doctor about it and things going on in your life." |
| Theme 2.2: Participants Experience Difficulty With Using the Apps | Many participants had difficulty using the apps, including patients who reported no prior difficulty with smartphone app usage. Translating apps into Spanish would enhance usability for Spanish-speaking participants | • "You would keep track of your blood pressure, you know that day it was a little high, now it's better. You would keep better track of your medication." |
| Theme 2.3: Participants Do Not Share their BP With Providers | Most participants did not currently share their BP. Some tracked BP, but did not share with their provider due to misplaced logs. | • "Really the main thing I liked [about the apps] is … how you can go on the app and you can get to your blood pressure feed. Just that fast... [You don't have to] go through a whole bunch of stuff. It's just right there. That's what I like about it." |
| | | • "[Logging in the app is] really good because every time I write it down, I always forget where I put it. I'm always losing it. It's great that my [BP recordings] stay [in the app]. I can just show it. I'd love that." |
| | | • "You meet your doctor once every 3 months, so if I take a reading now, I forget it after two weeks." |
| Analytical Code 3: Varied Patient Preferences on Home Blood Pressure Monitoring Device Features | | |
| Theme 3.1: Device Features Did Not Impact Overall Device Preference | Participants differed on preferences for BP cuff type and device size, but these features did not impact device preference. | • "I don't want anything gigantic. I want it to be perfectly small, where I can take it if I need be, to be able to take it with me if I'm traveling." |
| Theme 3.2: Participants Require Assistance With Bluetooth Pairing | Many participants needed help pairing the Bluetooth device, and some preferred the non-Bluetooth option because of this requirement | • "This is pretty big. So it won't be good for traveling too much." |
| Theme 3.3: Cost as a Deciding Factor in Device Preference | While participants preferred BP results to be automatically shared with their provider, they would not pay for this feature and also had concerns related to the cost of batteries. | • "Just sending a message – if it's going to cost you money – that's a rip off." |
| | | • "[Paying for remote patient monitoring] is a turn off. If I had to pay more money, I'd rather not." |
| | | • "If it can be sent to my [doctor] without me knowing, that'd be great… [But I don't want to be] paying for [that]." |
| | | • "The only thing holding me back from [this device] is paying." |
| | | • "People don't have access to batteries, like you or the manufacturer think. You think an old man of 70 - 80 [years of age] would go out to buy a battery?" |

[a]SMBP: Self-measured blood pressure monitoring

[b]BP: blood pressure.

## Analytical Code 1: Patient Knowledge About SMBP

### *Inconsistencies and Gaps Exist in Patient Education*

Many reported no formal training on using their BP monitor or normal BP range. Instead, many learned by observing providers in clinics, reading the manual, or watching online video tutorials. A few participants received demonstrations or were told about BP at clinic visits but did not remember what they learned and were not given additional materials.

Existing knowledge about BP was incomplete and varied. Few participants reported understanding the idea of normal BP range. Some participants noted the importance of taking repeated measurements, staying calm before a reading, and the importance of weight and diet for BP management.

### *Patients Desire Training and Education*

Nearly all participants wanted more training about BP devices and at-home measurement, preferring in-person demonstrations. Some said written instructions or video tutorials may suffice. In addition to training on BP device usage, some participants wanted information about BP ranges. Nearly all participants expected clinic support if they encountered an issue when measuring BP. Some participants also mentioned troubleshooting with family.

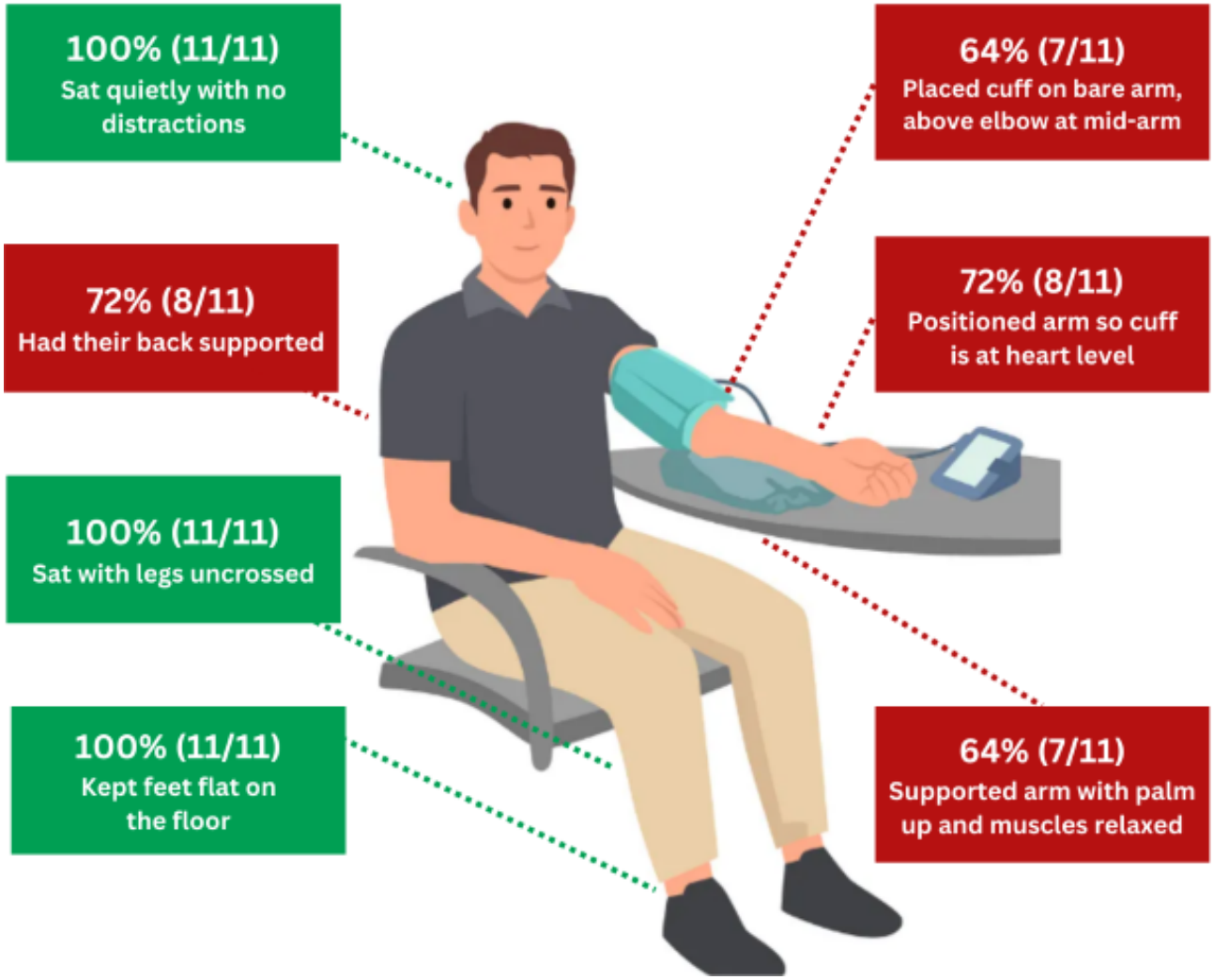
Participants valued having additional written and video resources. They preferred written resources that were easy to read and understand, larger with large and bolded font, and contained concise, numbered steps to follow. Some participants also thought visual illustrations would be helpful. Participants noted that the instruction manuals included with the BP monitoring devices were detailed and had a lot of information, which some felt was overwhelming. Spanish-speaking participants wanted to have Spanish materials.

### *Incorrect Technique and Missing Considerations in Home BP Measurement*

Most participants did not follow all guidelines set forth by the American Heart Association (AHA) and American Medical Association (AMA) for accurately measuring BP at home [28] and displayed incorrect technique when demonstrating an at-home BP measurement (Figure 1). Of the 11 participants who were asked to demonstrate how they would measure their BP at home, only 4 were consistent with all guidelines set forth by the AMA and AHA. The most common missed techniques in BP measurement were placing the cuff on the bare arm, above the elbow at midarm; ensuring the arm is supported, with palm up and muscles relaxed; positioning the arm at heart level; and ensuring the back is supported. All participants properly uncrossed their legs, rested their feet flatly on the floor, and sat quietly without distractions (though we provided the environment).

When asked about considerations when measuring BP at home, nearly all participants did not consider the time of day or the timing of medications, eating meals, smoking, drinking alcohol, or using the restroom in relation to their readings. Most participants, however, did report resting or relaxing prior to taking a measurement.

**Figure 1.** Demonstration of blood pressure measurement technique evaluated by American Medical Association and American Heart Association guidelines.




## Analytical Code 2: Patterns and Challenges in Self-Monitoring Blood Pressure

### *Participants Like the Idea of Tracking BP*

Most participants did not currently track their BP but cited it as important or wanted to do so. Some would track BP only under particular circumstances, such as when their BP is unusually high.

Participants liked that the apps linked and tracked their BP results from the device. Nearly all participants appreciated visual aids (eg, graphs) that show their BP history. Many participants liked immediately seeing their results when opening the app; some participants specified that Device 2's app was simpler and straightforward to use. Some also appreciated additional features of Device 1's app, such as the option of keeping a diary to write notes. Participants appreciated visuals, a straightforward interface that enables easy app usage and BP tracking, and personalized tracking capabilities.

### *Participants Experienced Difficulty With Using the Apps*

Difficulty in app usage varied. When surveyed, all participants who reported difficulty using a smartphone without assistance or installing apps also had difficulty navigating app features or required assistance upon observation. Some who reported no difficulty in these 2 survey questions also had difficulty or required assistance upon demonstration. Nearly all participants who had trouble navigating the apps were Spanish-speaking and expressed that translating apps entirely into Spanish would enhance ease of use.

### *Participants Do Not Share Their BP With Providers*

Furthermore, most participants did not currently share their BP with their provider. Most participants who track their BP often misplace their measurement logs. Two participants shared with their providers only if they remembered where they kept their readings.

## Analytical Code 3: Varied Patient Preferences on Home Blood Pressure Monitoring Device Features

### *Device Features Did Not Impact Overall Device Preference*

Participants prioritize comfort and ease of use when evaluating 2 different BP cuffs, but this preference was split between hard and soft cuffs. Preferences also differed between the smaller size of Device 2 and the larger display of Device 1. However, cuff type and portability of the devices did not impact overall device preference.

### *Participants Require Assistance With Bluetooth Pairing*

Many participants did not view Bluetooth pairing, a feature of Device 1, as a deterrent to measuring BP, with half of them citing familiarity with Bluetooth. We also observed that these participants successfully connected the device with Bluetooth. Nearly all other participants required help from study staff to pair the Bluetooth device. A few explicitly cited not wanting to deal with pairing or repairing Bluetooth. Some preferred Device 2 because it did not require Bluetooth or said that Device 1 was more difficult to use for this reason.

### *Cost as a Deciding Factor in Device Preference*

Moreover, while approximately half of the participants value having their BP results automatically shared with their provider (such as would occur in a cellular-enabled monitor like Device 2), nearly all would not pay for this feature. One was open to having results automatically sent if in poor health. A few would pay US $5-US $15 a month if required, but they were strongly opposed. If this feature were free, three participants would prefer Device 2. Most strongly preferred having a plug-in charging option for the battery-operated devices because of concerns related to accessing or purchasing future batteries.

## Discussion

### Principal Findings

In our qualitative study of safety net patients with hypertension, the cost of device features, gaps in existing BP knowledge, and lack of training and resources presented challenges to SMBP adoption. Our findings highlight the need to provide affordable, language-concordant resources and comprehensive training to leverage SMBP for hypertension management in safety net patients.

Our findings of participant motivation for SMBP and barriers related to cost and health literacy among low-income and minority communities were overall consistent with prior studies [11,14-16]. Our participants valued tracking their BP and are motivated to engage in SMBP practices and share results with their providers [11]. Moreover, participants preferred having their readings automatically shared with their provider without requiring pairing to their own device, citing it as extremely helpful. However, nearly all participants were not willing to pay for this feature. Plug-in devices were also strongly preferred to avoid the cost of batteries. These findings suggest that SMBP adoption in lower-income and uninsured and Medicaid populations is impacted by affordability or payor coverage of SMBP devices that meet patients' needs. In addition, nearly all patients demonstrated inconsistencies in BP education, displayed incorrect BP measurement technique, and had not received formal training. These findings reinforce the notion that barriers to successful SMBP adoption stem from external care factors, such as cost and gaps in available training and resources, rather than patient motivation. Providers should evaluate patients' barriers and preferences when recommending SMBP.

In addition to cost, digital accessibility and literacy should be assessed as contributors to SMBP non-adoption. Aligned with studies that demonstrate socioeconomic status and Medicaid insurance as risk factors for digital exclusion [29], 8 of 15 participants exhibited some form of digital inaccessibility: lacking internet access, not activating their patient portal, or having difficulty connecting a device to Wi-Fi. Furthermore, most participants required assistance with pairing Bluetooth to Device 1 and with navigating app features on both devices, potentially indicating limited digital literacy. SMBP interventions should be complemented with patient training and resources. Importantly, it appears that cellular SMBP devices that automatically transmit BP readings to reduce digital complexity may promote sharing results with providers, and this should be considered as a focus of future research and implementation.

Moreover, Spanish-speaking participants reported that using apps was challenging because they were presented in English. This further supports the unmet need to have user-friendly, language-concordant digital SMBP tools [11,30]. Spanish-speaking participants also wanted manuals and training to be delivered in Spanish. Addressing language nonconcordance in training and resources for other prominent but less prevalent languages (eg, Arabic) [31] in addition to Spanish may further promote widespread SMBP adoption. Notably, patients who face barriers in digital literacy and language discordance may be especially vulnerable to SMBP adoption challenges.

While a previous study on home BP monitoring enrolled both English and Spanish speakers [15], our study was the first to assess patient preferences between 2 SMBP devices and the impact of language on SMBP practices. Our results further add to knowledge about patient preferences for communication modality of their BP results with their care team and barriers to SMBP among racially diverse, low-income populations.

### Limitations

Our study was limited by a small, convenient sample. However, our sample size aligns with findings from a recent systematic review suggesting that saturation in qualitative research can often be achieved within a narrow sample size of 9-17 interviews [18]. Participants were assessed at only 1 time point; SMBP skills and preferences may differ in real care scenarios. In addition, despite our efforts to ensure credibility through independent coding, team discussions, and participant validation of findings, researcher interpretation may still reflect inherent subject biases common to qualitative research. While we maintained reflexivity through field notes and team discussions about our positionality, as with all qualitative work, the lived experiences and perspectives of the study team, which included individuals with expertise in primary care, health equity, and public health, may have introduced implicit biases in theme identification. Although we described our study setting and participants in detail to support transferability, our single-site study within an urban safety net context may limit applicability to different health care settings. Future studies should explore how support systems (eg, family and caregivers) could impact SMBP adoption. Collaborating directly with patients to refine analytical domains and themes could also improve the transferability of our findings. An ongoing randomized controlled trial is currently underway to longitudinally assess device implementation and BP outcomes, which will help address these limitations [25].

## Conclusions

Patients' values and barriers can inform solutions that facilitate and improve patient self-management of hypertension. Our findings reinforce the importance of affordability, accessibility, and providing robust resources when implementing SMBP in diverse, safety net populations. The effectiveness of cellular-enabled SMBP devices should be further evaluated.



## Acknowledgments

We would like to acknowledge José Miramontes for completing a participant interview.

This project is supported by the following grants: K23HL157750, R18HS029817, R01HL159372.


## Data Availability

The data collected during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

JJS, VEK, MM, and ECK contributed to conceptualization. JJS, VEK, IL, CHK, FG, and CG performed data curation. JJS and VEK conducted analysis. JJS and VEK contributed to drafting manuscript

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide in English.
[DOCX File, 67 KB - cardio_v9i1e60196_app1.docx ]

Multimedia Appendix 2
Interview guide in Spanish.
[DOCX File, 51 KB - cardio_v9i1e60196_app2.docx ]

## References


1. Park B, Budzynska K, Almasri N, et al. Tight versus standard blood pressure control on the incidence of myocardial infarction and stroke: an observational retrospective cohort study in the general ambulatory setting. BMC Fam Pract 2020 May 16;21(1):91. [doi: 10.1186/s12875-020-01163-4] [Medline: 32416722]
2. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart disease and stroke statistics-2023 update: a report from the American Heart Association. Circulation 2023 Feb 21;147(8):e93-e621. [doi: 10.1161/CIR.0000000000001123] [Medline: 36695182]
3. Fuchs FD, Whelton PK. High blood pressure and cardiovascular disease. Hypertension 2020 Feb;75(2):285-292. [doi: 10.1161/HYPERTENSIONAHA.119.14240] [Medline: 31865786]
4. Uhlig K, Patel K, Ip S, Kitsios GD, Balk EM. Self-measured blood pressure monitoring in the management of hypertension: a systematic review and meta-analysis. Ann Intern Med 2013 Aug 6;159(3):185-194. [doi: 10.7326/0003-4819-159-3-201308060-00008] [Medline: 23922064]
5. Cappuccio FP, Kerry SM, Forbes L, Donald A. Blood pressure control by home monitoring: meta-analysis of randomised trials. BMJ 2004 Jul 17;329(7458):145. [doi: 10.1136/bmj.38121.684410.AE] [Medline: 15194600]
6. Shimbo D, Artinian NT, Basile JN, et al. Self-measured blood pressure monitoring at home: a joint policy statement from the American Heart Association and American Medical Association. Circulation 2020 Jul 28;142(4):e42-e63. [doi: 10.1161/CIR.0000000000000803] [Medline: 32567342]
7. Angell S, Guthartz S, Dalal M, et al. Integrating self blood pressure monitoring into the routine management of uncontrolled hypertension: translating evidence to practice. J of Clinical Hypertension 2013 Mar;15(3):180-185 [FREE Full text] [doi: 10.1111/jch.12046]
8. Muntner P, Hardy ST, Fine LJ, et al. Trends in blood pressure control among US adults with hypertension, 1999-2000 to 2017-2018. JAMA 2020 Sep 22;324(12):1190-1200. [doi: 10.1001/jama.2020.14545] [Medline: 32902588]
9. Shahu A, Herrin J, Dhruva SS, et al. Disparities in socioeconomic context and association with blood pressure control and cardiovascular outcomes in ALLHAT. J Am Heart Assoc 2019 Aug 6;8(15):e012277. [doi: 10.1161/JAHA.119.012277] [Medline: 31362591]
10. Boulware LE, Ephraim PL, Hill-Briggs F, et al. Hypertension self-management in socially disadvantaged African Americans: the Achieving Blood Pressure Control Together (ACT) randomized comparative effectiveness trial. J Gen Intern Med 2020 Jan;35(1):142-152. [doi: 10.1007/s11606-019-05396-7] [Medline: 31705466]

11. Khoong EC, Commodore-Mensah Y, Lyles CR, Fontil V. Use of self-measured blood pressure monitoring to improve hypertension. Curr Hypertens Rep 2022 Nov;24(11):599-613. [doi: 10.1007/s11906-022-01218-0] [Medline: 36001268]
12. Fang J, Yang Q, Ayala C, Loustalot F. Disparities in access to care among US adults with self-reported hypertension. Am J Hypertens 2014 Nov;27(11):1377-1386. [doi: 10.1093/ajh/hpu061] [Medline: 24847953]
13. Kim EJ, Kim T, Paasche-Orlow MK, Rose AJ, Hanchate AD. Disparities in hypertension associated with limited English proficiency. J Gen Intern Med 2017 Jun;32(6):632-639. [doi: 10.1007/s11606-017-3999-9] [Medline: 28160188]
14. Parmar S, Commodore-Mensah Y, Singh G, et al. Patient practices, perceptions, and barriers to self-measurement of blood pressure in a rural health systemAsk ChatGPT. Am J Hypertens 2024 Oct 14;37(11):876-883. [doi: 10.1093/ajh/hpae085] [Medline: 38932512]
15. Carter EJ, Moise N, Alcántara C, Sullivan AM, Kronish IM. Patient barriers and facilitators to ambulatory and home blood pressure monitoring: a qualitative study. Am J Hypertens 2018 Jul 16;31(8):919-927. [doi: 10.1093/ajh/hpy062] [Medline: 29788130]
16. Bridi L, Albahsahli B, Bencheikh N, et al. Barriers and facilitators to self-measured blood pressure monitoring among US-resettled Arab refugees with hypertension: a qualitative study. BMC Prim Care 2023 Nov 30;24(1):256. [doi: 10.1186/s12875-023-02215-1] [Medline: 38036967]
17. America's Health Care Safety Net: Intact but Endangered: National Academies Press; 2000:9612. [doi: 10.17226/9612]
18. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523. [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
19. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
20. St George SM, Harkness AR, Rodriguez-Diaz CE, Weinstein ER, Pavia V, Hamilton AB. Applying rapid qualitative analysis for health equity: lessons learned using "EARS" with Latino communities. Int J Qual Methods 2023;22:16094069231164938. [doi: 10.1177/16094069231164938] [Medline: 38463016]
21. Neal JW, Neal ZP, VanDyke E, Kornbluh M. Expediting the analysis of qualitative data in evaluation: a procedure for the Rapid Identification of Themes From Audio Recordings (RITA). Am J Eval 2015 Mar;36(1):118-132. [doi: 10.1177/1098214014536601]
22. Nevedal AL, Reardon CM, Opra Widerquist MA, et al. Rapid versus traditional qualitative analysis using the Consolidated Framework for Implementation Research (CFIR). Implement Sci 2021 Jul 2;16(1):67. [doi: 10.1186/s13012-021-01111-5] [Medline: 34215286]
23. Lewinski AA, Crowley MJ, Miller C, et al. Applied rapid qualitative analysis to develop a contextually appropriate intervention and increase the likelihood of uptake. Med Care 2021 Jun 1;59(Suppl 3):S242-S251. [doi: 10.1097/MLR.0000000000001553] [Medline: 33976073]
24. Comparing ways to support patients with home blood pressure monitoring -- the CHARMED study. Pcori. 2023. URL: https://www.pcori.org/research-results/2023/comparing-ways-support-patients-home-blood-pressure-monitoring-charmed-study [accessed 2025-02-08]
25. Fontil V, Khoong EC, Green BB, et al. Randomized trial protocol for remote monitoring for equity in advancing the control of hypertension in safety net systems (REACH-SNS) study. Contemp Clin Trials 2023 Mar;126:107112. [doi: 10.1016/j.cct.2023.107112] [Medline: 36738916]
26. Implementation strategies for self-measured blood pressure monitoring in racially and ethnically diverse populations (ins2pired). RePORTER. URL: https://reporter.nih.gov/search/wELMagNM002aa5qSTAcvgg/project-details/10831792#details [accessed 2025-02-08]
27. Korstjens I, Moser A. Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. Eur J Gen Pract 2018 Dec;24(1):120-124. [doi: 10.1080/13814788.2017.1375092] [Medline: 29202616]
28. American Medical Association. How to Measure Your Blood Pressure at Home: American Medical Association; 2020. URL: https://www.heart.org/-/media/Files/Health-Topics/High-Blood-Pressure/measuringbpathome.pdf
29. Lyles CR, Nguyen OK, Khoong EC, Aguilera A, Sarkar U. Multilevel determinants of digital health equity: a literature synthesis to advance the field. Annu Rev Public Health 2023 Apr 3;44(383–405):383-405. [doi: 10.1146/annurev-publhealth-071521-023913] [Medline: 36525960]
30. Sarkar U, Lisker S, Lyles CR. How to narrow the digital divide in US health care. Harvard Buisness Review. URL: https://hbr.org/2021/05/how-to-narrow-the-digital-divide-in-u-s-health-care [accessed 2025-07-21]
31. Dietrich S, Hernandez E. What languages do we speak in the united states? United States Census Bureau. 2022. URL: https://www.census.gov/library/stories/2022/12/languages-we-speak-in-united-states.html [accessed 2025-07-21]


## Abbreviations

**AHA:** American Heart Association
**AMA:** American Medical Association
**BP:** blood pressure
**SMBP:** self-measured blood pressure monitoring

*Edited by A Coristine; submitted 04.05.24; peer-reviewed by M Wolff, P Nymberg; revised version received 27.06.25; accepted 29.06.25; published 29.08.25.*

*Please cite as:*
*Shih JJ, Kwok VE, Luna I, Kim HC, Garcia F, Gutierrez C, Garcia M, Lyles CR, Khoong EC*
*English- and Spanish-Speaking Patient Preferences on Home Blood Pressure Monitors in an Urban Safety Net Setting: Qualitative Study*
*JMIR Cardio 2025;9:e60196*
*URL: https://cardio.jmir.org/2025/1/e60196*
*doi:10.2196/60196*

# Causal Inference for Hypertension Prediction With Wearable Electrocardiogram and Photoplethysmogram Signals: Feasibility Study

Ke Gong, MEng; Yifan Chen, PhD; Xinyue Song, MEng; Zhizhong Fu, MEng; Xiaorong Ding, PhD

School of Life Science and Technology, University of Electronic Science and Technology of China, Research Building C348A, 3rd Fl, Chengdu, China

**Corresponding Author:**
Xiaorong Ding, PhD
School of Life Science and Technology, University of Electronic Science and Technology of China, Research Building C348A, 3rd Fl, Chengdu, China

## Abstract

**Background:** Hypertension is a leading cause of cardiovascular disease and premature death worldwide, and it puts a heavy burden on the health care system. Therefore, it is very important to detect and evaluate hypertension and related cardiovascular events to enable early prevention, detection, and management. Hypertension can be detected in a timely manner with cardiac signals, such as through an electrocardiogram (ECG) and photoplethysmogram (PPG), which can be observed via wearable sensors. Most previous studies predicted hypertension from ECG and PPG signals with extracted features that are correlated with hypertension. However, correlation is sometimes unreliable and may be affected by confounding factors.

**Objective:** The aim of this study was to investigate the feasibility of predicting the risk of hypertension by exploring features that are causally related to hypertension via causal inference methods. Additionally, we paid special attention to and verified the reliability and effectiveness of causality compared to correlation.

**Methods:** We used a large public dataset from the Aurora Project, which was conducted by Microsoft Research. The dataset included diverse individuals who were balanced in terms of gender, age, and the condition of hypertension, with their ECG and PPG signals simultaneously acquired with wrist-worn wearable devices. We first extracted 205 features from the ECG and PPG signals, calculated 6 statistical metrics for these 205 features, and selected some valuable features out of the 205 features under each statistical metric. Then, 6 causal graphs of the selected features for each kind of statistical metric and hypertension were constructed with the equivalent greedy search algorithm. We further fused the 6 causal graphs into 1 causal graph and identified features that were causally related to hypertension from the causal graph. Finally, we used these features to detect hypertension via machine learning algorithms.

**Results:** We validated the proposed method on 405 subjects. We identified 24 causal features that were associated with hypertension. The causal features could detect hypertension with an accuracy of 89%, precision of 92%, and recall of 82%, which outperformed detection with correlation features (accuracy of 85%, precision of 88%, and recall of 77%).

**Conclusions:** The results indicated that the causal inference–based approach can potentially clarify the mechanism of hypertension detection with noninvasive signals and effectively detect hypertension. It also revealed that causality can be more reliable and effective than correlation for hypertension detection and other application scenarios.

***(JMIR Cardio 2025;9:e60238)*** doi:10.2196/60238

**KEYWORDS**

hypertension; causal inference; wearable physiological signals; electrocardiogram; photoplethysmogram

## Introduction

Hypertension, also known as high blood pressure (BP), is a condition in which the pressure of the blood increases in the arteries. The diagnosis of hypertension relies on BP measurement, and it is defined as systolic BP (SBP) ≥140 mm Hg or diastolic BP (DBP) ≥90 mm Hg [1]. Hypertension can be further classified into 3 stages. Stage 1 hypertension is associated with SBP and DBP ranges of 140 - 159 mm Hg and 90 - 99 mm Hg, respectively. Stage 2 hypertension is characterized by SBP and DBP ranges of 160 - 179 mm Hg and 100 - 109 mm Hg, respectively. For stage 3 hypertension, the SBP and DBP are more than 180 mm Hg and 110 mm Hg [1,2].

Furthermore, it is noteworthy that even when SBP ≥115 mm Hg and DBP ≥75 mm Hg, a continuous relationship exists between the increase in BP level and the occurrence of cardiovascular or renal pathological conditions and even fatal events. The definition of high blood pressure as SBP ≥140 mm Hg or DBP ≥90 mm Hg primarily serves the purpose of

simplifying hypertension diagnosis and decision-making regarding hypertension treatment. This threshold was chosen because the benefits of intervention outweigh the risks associated with nonintervention in this context.

According to a review of the global epidemiology of hypertension [3], hypertension is a leading preventable risk factor for cardiovascular disease and all-cause mortality worldwide. In 2010, a total of 1.38 billion people (31.1% of the global adult population) had hypertension. The prevalence of hypertension is rising globally owing to the aging of the population and increases in exposure to lifestyle risk factors, including unhealthy diets and lack of physical activity.

In addition, hypertension can be divided into primary and secondary forms. Secondary hypertension originates from specific causes and only encompasses a small fraction of the population. Primary hypertension covers the remaining large fraction of the hypertension population, and it arises from intricate interactions among genetic factors, environmental influences, and the aging process. These factors collectively contribute to an increase in systemic vascular resistance, a hallmark hemodynamic abnormality that leads to elevated BP in almost all hypertensive individuals [4]. Furthermore, considering that hypertension may not show any symptoms in its early stages and that there is a continuous relationship between an increase in BP and the risk of stroke, coronary heart disease, heart failure, and chronic kidney disease, it is very important to detect and treat hypertension in the early stages.

Moreover, physicians often diagnose hypertension by office BP, but masked hypertension and white coat hypertension cannot be effectively detected by office BP. Instead, they usually detect masked hypertension and white coat hypertension through a 24-hour ambulatory recording of the BP signal [5], but this process is cumbersome. Hence, there are data-driven approaches based on noninvasive signals for the detection of hypertension, such as electrocardiogram (ECG) or photoplethysmogram (PPG), that are easily accessible from wearable sensors [2]. Subsequently, wearable monitoring can continuously monitor patients' physiological conditions 24 hours a day. Compared with outpatient blood pressure monitoring, wearable monitoring can obtain patients' rhythm information and real physiological conditions (to avoid white coat hypertension and other conditions), as well as the impact of patients' behaviors on physiological indicators and other personalized information. Rich reference information is conducive to more accurate assessment and stratification of individual risks.

There are various studies on detecting hypertension with data-driven methods based on noninvasive signals. These methods include classic machine learning models with hand-extracted features and feature representation learning with deep learning methods. For example, Paragliola et al [6] proposed a novel approach for analyzing and classifying the ECG signal with a hybrid deep learning network method called hybrid deep network, which combines long short-term memory, convolutional neural networks, and deep neural networks. The hybrid method can reach an average accuracy of 0.98 and an average sensitivity and specificity of 0.97. Elgendi et al [7] reviewed the effect of different types of artifacts added to the PPG signal, characteristic features of the PPG waveform, and existing indexes on hypertension diagnosis. In another study, Alkhodari et al [8] used features related to heart rate variability to predict hypertension based on decision trees and random undersampling boosting. The accuracy of the method was 0.81, with the $F_1$-score and area under the receiver operating characteristic curve (AUC) being 0.86 and 0.89, respectively. In a study about the automated detection of hypertension severity, Rajput et al [9] developed a 2-band optimal orthogonal wavelet filter bank method, which generates 6 subbands from each ECG signal through a 5-level wavelet decomposition. Further, the sample mean and wavelet entropy features of all subbands were computed to predict the risk of hypertension with classic machine learning methods, such as k-nearest neighbors and support vector machine, and the proposed method can achieve an average classification accuracy of 0.99.

However, most of the previously mentioned studies relied on extracting features correlated with hypertension but ignored the causality of hypertension and characteristic variables. Due to the presence of confounding factors, correlations can lead to wrong conclusions, just like Simpson's paradox [10]. In different populations, the distribution of confounding factors will change, which means the correlations can be unstable and unreliable. Instead, causal inference can not only identify more reliable feature variables with the elimination of confounding factors but also provide more trustworthy guidance for further exploring the physiological mechanisms of hypertension [11].

In this work, we propose to predict hypertension based on causal inference with wearable noninvasive signals. The overview of the proposed method is delineated in Figures 1 and 2. We will select effective features based on causality between hypertension and features extracted from PPG and ECG signals. Then, combined with the detected causal features, we will predict hypertension and evaluate its prediction performance by various evaluation metrics. Ultimately, we aim to identify some features that may be of great value in predicting hypertension.

**Figure 1.** Research route flow chart.

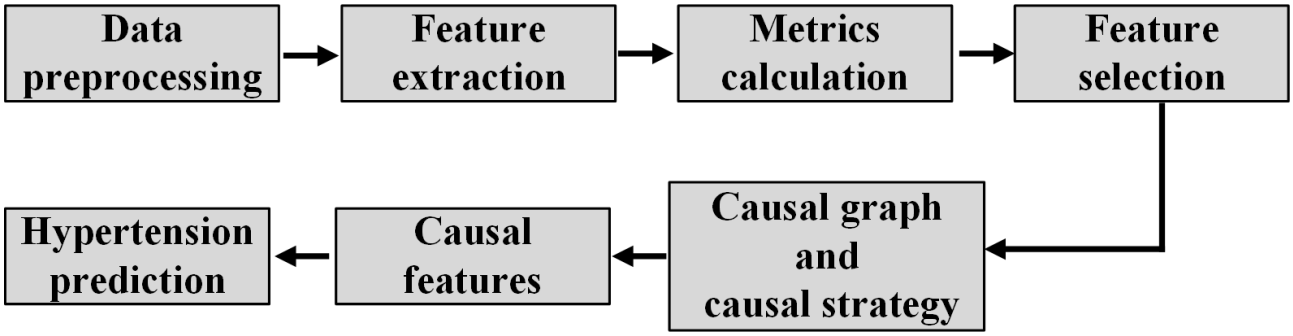
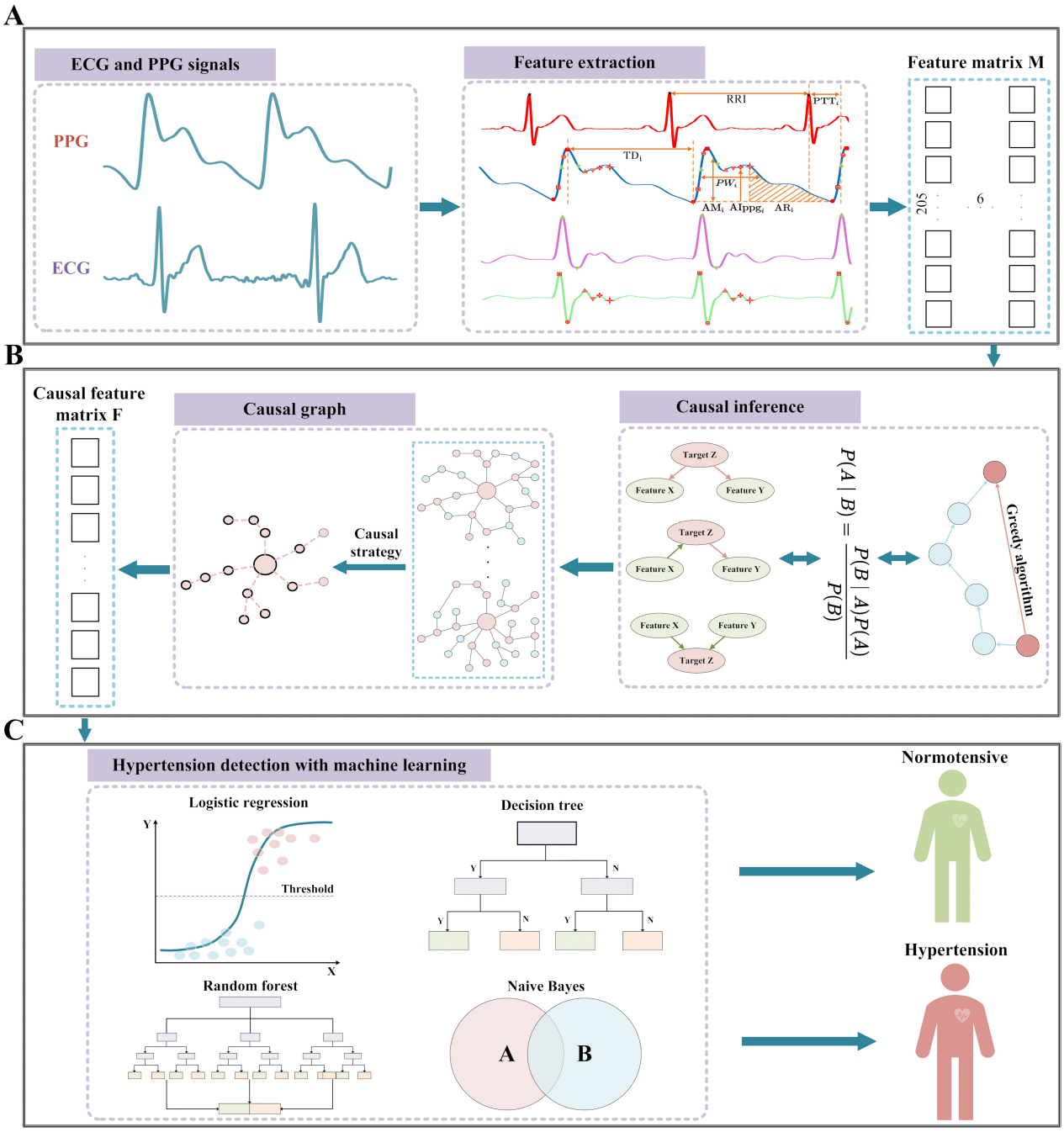


**Figure 2.** Flowchart of the causal inference for hypertension prediction. (A) Signal preprocessing: 205-dimension beat-by-beat features were extracted from the ECG and PPG as well as the first and second derivatives of the PPG signal (dPPG, sdPPG), and the statistical metrics of these features were calculated as the feature matrix M. (B) Based on the feature matrix M, the causal graphs of the extracted features and hypertension status were identified with the causal inference algorithm (the equivalent greedy search algorithm). (C) The causal feature matrix F was identified from the causal graph obtained from step (B), and we used machine learning classification algorithms to achieve hypertension prediction. ECG: electrocardiogram; PPG: photoplethysmogram.




## Methods

The methods of this paper can be divided into 7 steps; the details of each step are shown in Figure 1.

### Ethical Considerations

In this study, we used data from the Microsoft Waveform Database, and we obtained data access permission from the Microsoft Data Access Committee [12]. Microsoft obtained institutional review board approval from WCG IRB (Puyallup, WA, United States). Individuals unable to consent in English, pregnant women, prisoners, institutionalized individuals, and individuals younger than 18 years were excluded from participation due to their vulnerable status. All the subjects voluntarily participated in the experiment and signed informed consent. The original informed consent and the institutional review board both allow for secondary analysis without additional consent. The dataset used in this study was de-identified to protect the privacy of the subjects.

### Data

The database that we obtained data from was developed for validating new methods for blood pressure measurement with

noninvasive sensors. Noninvasive epidermal pressure signals, ECG signals, and PPG signals were acquired with tension, electrical, and optical sensors, respectively. Meanwhile, the reference blood pressure was measured with either the oscillometric method or the auscultatory method. In this study, we used noninvasive signals for hypertension detection. To validate our proposed method, we used data collected based on the oscillometric method. A total of 614 subjects participated in the oscillometric protocol scheme, with ages ranging from 18-85 years. After excluding data anomalies during the collection process, including miswear, malfunction, data file failure, participant opt-out, alignment failure, and quality failure, relevant measurement information from 483 subjects was retained [12]. In a further waveform preprocessing step, poor waveform segments and subjects with less than 4 qualified waveform segments were removed, which led to the final retention of measurement data from 405 participants, comprising 183 hypertensive patients and 222 healthy individuals. The ages of the 405 participants ranged from 18-60 years, with an average age of 45 years. In addition, the 405 participants comprised 199 females and 206 males.

Moreover, measurements in this protocol were obtained during controlled laboratory visits spaced at least 24 hours apart. Additionally, dynamic measurements were collected during the 24-hour interval between laboratory visits. Automatic measurements were taken every 30 minutes in the morning and every 60 minutes in the evening. Each patient typically had 24-36 waveform segments, with each acquired for 15-30 seconds. Our feature extraction primarily relied on data from dynamic measurements.

## Feature Extraction

We extracted 205 features from the filtered ECG and PPG signals with the extraction method defined in our previous study [13]. The features mainly include pulse transit time (PTT), time duration (TD), amplitude (AM), intensity of PPG, the first derivative of PPG (dPPG), the second derivative of PPG (sdPPG), area under the PPG curve (AR), and physiological meaningful relative index (RI). The mathematical expression and definition of these features are as follows and are also described in Table 1. The fiducial points of ECG, PPG, dPPG, and sdPPG signals of each cardiac cycle were identified to calculate the features. The identified fiducial points are illustrated in Figure 3.

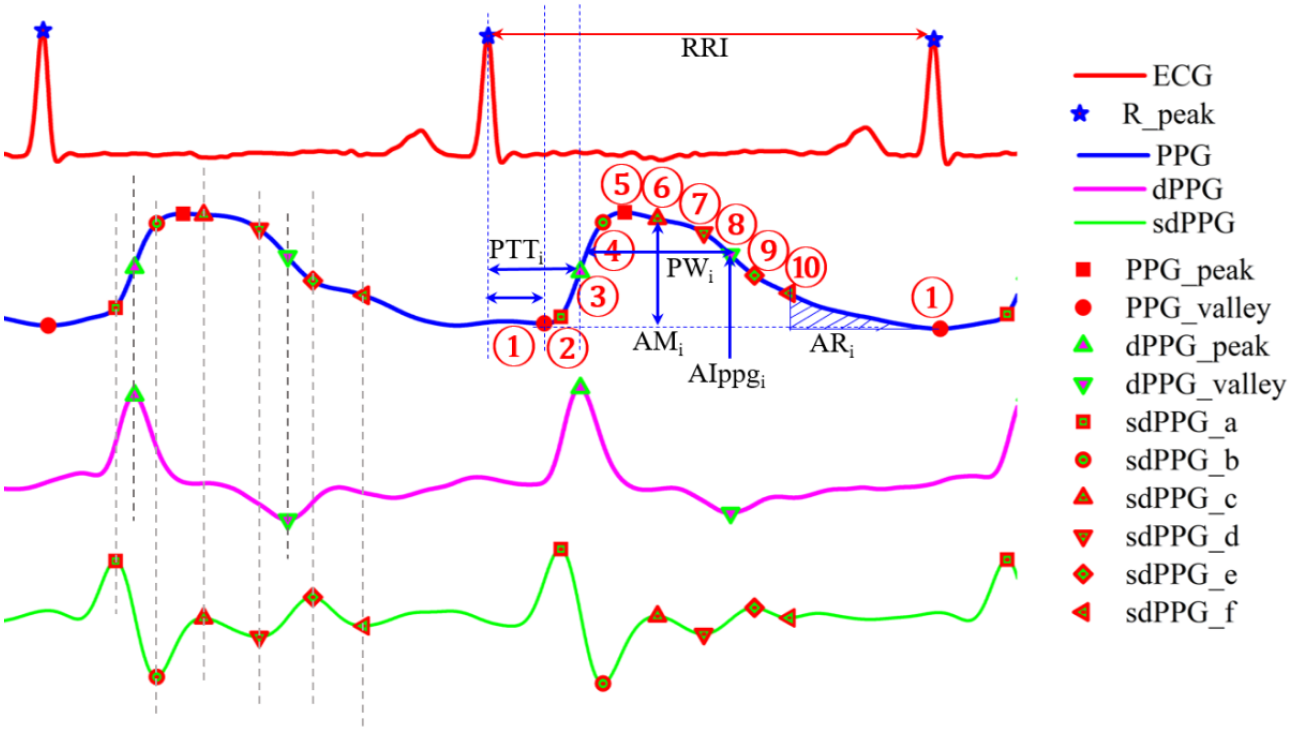
**Table .** Features extracted from electrocardiogram and photoplethysmogram signals.

| Index | Classification | Definition of features |
|---|---|---|
| 1 - 10 | Pulse transit time | Time deviation between R peak of electrocardiogram and fiducial points of photoplethysmogram |
| 11 - 66 | Time duration | Time duration between 2 fiducial points of photoplethysmogram |
| 67 - 111 | Amplitude | Amplitude between fiducial points of photoplethysmogram |
| 112 - 130 | Pulse intensity | Intensity of photoplethysmogram, dPPG[a], and sdPPG[b] at fiducial points |
| 131 - 185 | Area | Area under the photoplethysmogram curve between fiducial points |
| 186 - 205 | Relative index | Physiological meaningful ratio index |

[a]dPPG: the first derivative of photoplethysmogram.

[b]sdPPG: the second derivative of photoplethysmogram.

**Figure 3.** Diagram of fiducial points of the ECG and PPG signals as well as major types of features [13]. AI: absolute intensity; AR: area under the PPG curve; AM: amplitude; dPPG: the first derivative of PPG; ECG: electrocardiogram; PPG: photoplethysmogram; PTT: pulse transit time; PW: pulse width; RRI: R-R interval; sdPPG: the second derivative of PPG.




*Feature Point (FP, 1~10)* = [PPG valley, sdPPG a, dPPG peak, sdPPG a, PPG peak, sdPPG c, sdPPG d, dPPG valley, sdPPG e, sdPPG f, PPG valley next]

*PTT* = FP(i) - R peak, i=1~10

*TD* = [RRI, (FP(j) - FP(i)), i,j=1~10, and $j>i$]

*AM* = PPG(FP(j)) - PPG(FP(i)), i=1~10, and $j>i$

*AIPPG* = PPG(FP(i)), i=1~10

*AIdPPG* = dPPG(FP(i)), i=1~10

*AIsdPPG* = sdPPG(FP(i)), i=2,4,7~10

*AR* = Area between (FP(j) – FP(i)), i,j=1~10

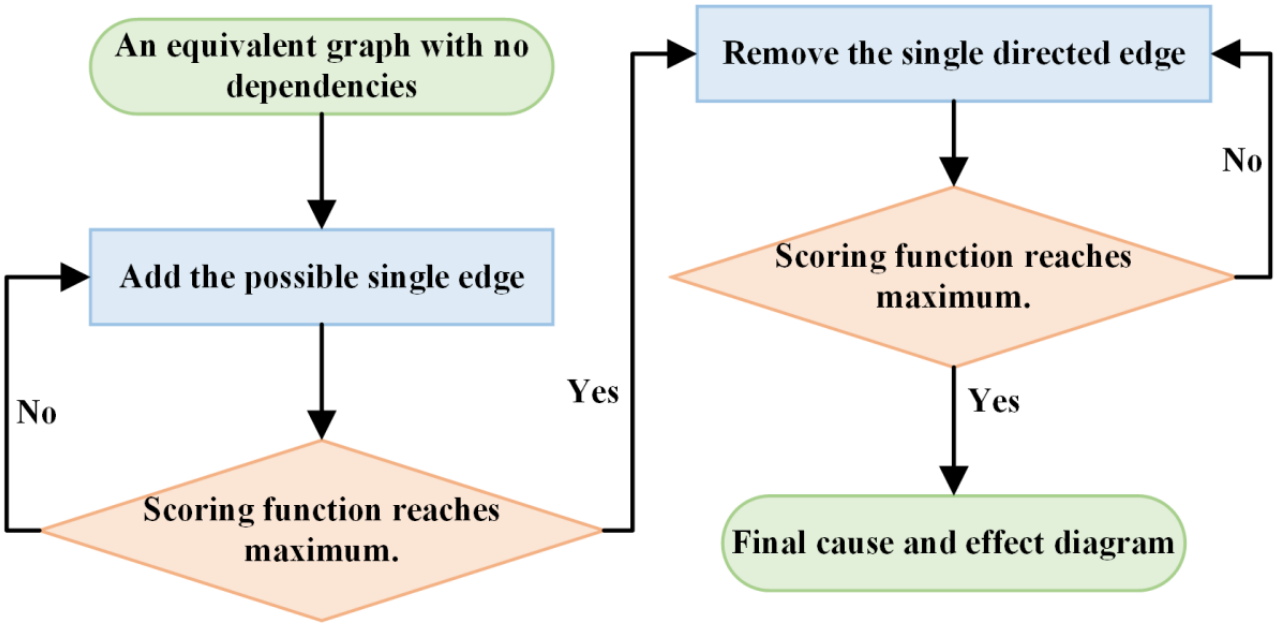
*RI*: relative rising time, dicrotic diastolic ratio, augmentation index, inflection point area point, slope transit time, ratio of sdPPG (b/a, c/a, (c+d–b)/a, etc), PPG intensity ratio, perfusion index [13].

After obtaining the above features, we can perform feature selection and build a causal graph based on the causal inference algorithm.

## Algorithm of Causal Inference

We used the greedy equivalence search (GES) algorithm to learn the causal graph. The GES algorithm is based on the theoretical basis of Meek's conjecture [14]. The Meek's conjecture is: if direct acyclic graph (DAG) M is an independent map of another DAG F, then there exists a finite set of edges in DAG F that can be added or reversed, after each modifiable edge is added or reversed direction, DAG M is still an independent graph of DAG F. After all modifications are done, M = F. Underlying the Meek's conjecture, we can use generalized score functions [15] and the GES algorithm to get the final causal graph. Figure 4 shows the implementation steps of the GES algorithm. In addition, we also provide the pseudo code to illustrate the detailed steps of the GES algorithm as shown in Textbox 1.

**Figure 4.** Flowchart of the greedy equivalence search algorithm.




**Textbox 1.** Algorithm 1: Apply-edge-operation(G,H).

Input: DAGsG and H where G≤H and G≠H

1: Set G′←G

2: While G and H contain a node Y that is a sink node in both DAGs and for which PaYG=PaYH, remove Y and all incident edges from both DAGs

3: end while

4: Let Y be any sink node in H

5: if Y has no children in G then

6: Let X be any parent of Y in H that is not a parent of

7: Y in G, add the edge X→Y

8: return G′

9: end if

10: Let DeYG denote the descendants of Y inG

11: And let D ∈ DeY G denote the (unique) maximal element from this set within 2

12: Let Z be any maximal child of Y in G such that G is a descendant of Y in G

13: if Y→Z is covered in G

14: reverse Y→Z in G′

15: Return G′

16: end if

17: if There exists a node X that is a parent of Y but not a parent of Z in G′ then

18: add X→Z to G′

19: return G′

20: end if

21: Let X be any parent of Z that is not a parent of Y

22: Add Y→Y to G′

23: return G′

Output: DAG G′ that results from adding or reversing an edge in G.

Then, the GES algorithm has 2 stages. In the first stage, it starts from an equivalence class (empty graph) with no dependencies and keeps adding possible edges to search for the largest equivalence class of generalized scoring functions until the

scoring functions' local maximum is reached. Then, in the second stage, the greedy principle is used to gradually delete the directed edges until the generalized scoring function reaches the local maximum again, and the final causal graph is obtained.

Considering that hypertension is a discrete variable while the feature variables are continuous, we are essentially dealing with mixed data. Traditional scoring functions such as Bayesian information criterion and Bayesian Dirichlet equivalent uniform do not take into account the issue of mixed data; for example, it discretizes continuous data and process it uniformly, resulting in a loss of valuable information. Therefore, we introduce a generalized scoring function to replace traditional scoring functions. The generalized function is primarily based on kernels and handles linear causal relationships, nonlinear causal relationships, continuous variables, discrete variables, and mixed data in a uniform manner, maximizing information retention. Finally, this scoring function addresses the issue of Markov equivalence classes, to some extent, overcoming the limitation of equivalence greedy search algorithms in distinguishing Markov equivalence classes.

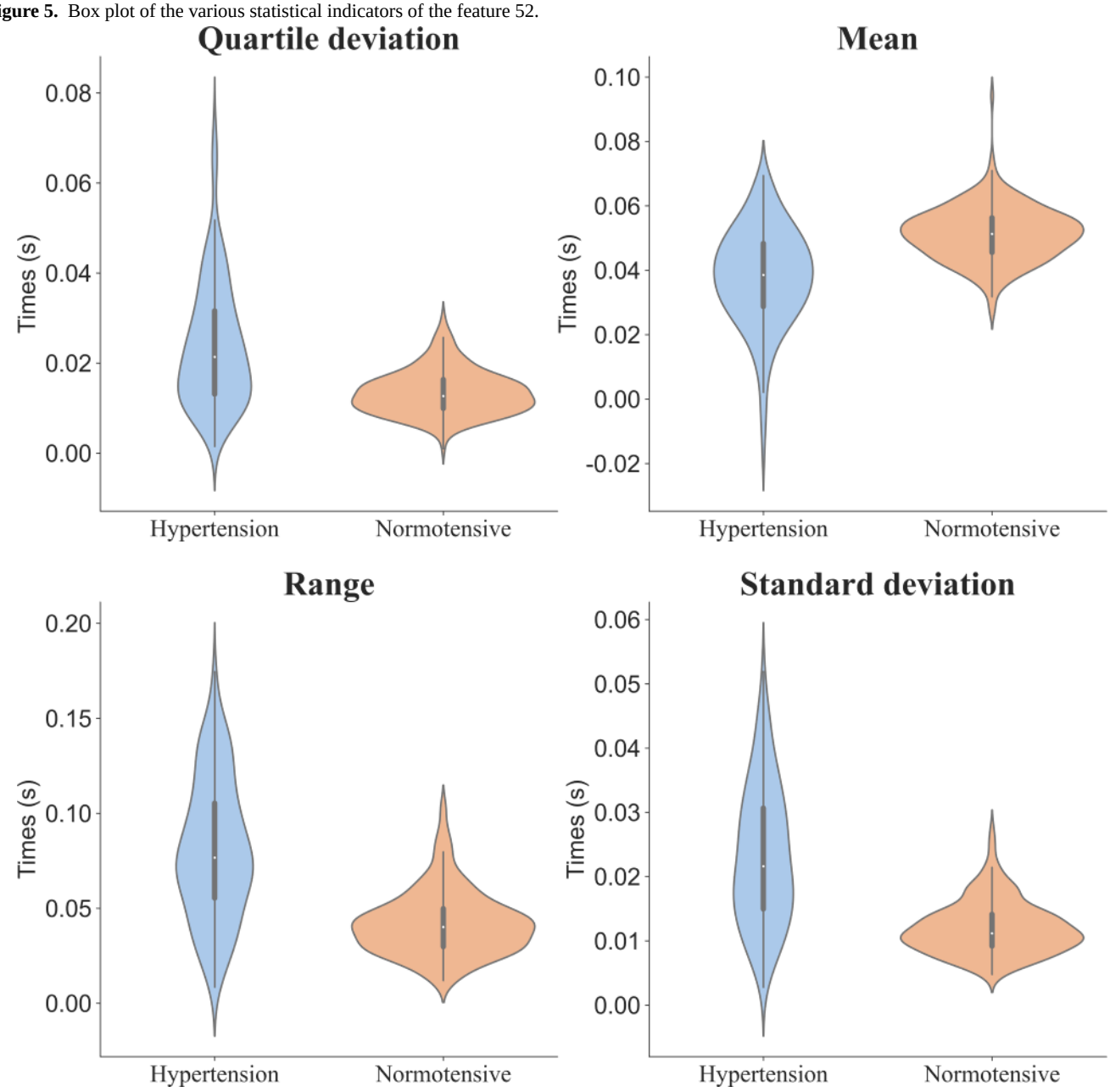
Finally, we needed to organize a feature matrix in which each row represents a sample and each column represents a kind of feature, then input this matrix into the equivalent greedy search algorithm to obtain the causal graph. Prior to this, feature selection is a necessary step to construct the feature matrix.

### Feature Selection

This section mainly explains the specific process of feature selection in this study, which is mainly divided into the following 3 parts. After completing feature selection, we will perform causal strategy and causal graph construction.

1. Six statistical metrics: Since ECG and PPG signals are time series data, we extracted the beat-by-beat features and calculated the statistical metrics of these 205 features to represent the temporal variability information. The statistical metrics include: standard deviation, range, mean, quartile deviation, coefficient of variation, and median, which result in 205×6=1230 dimensional features. This allows us to capture and analyze the temporal characteristics of ECG and PPG signals while summarizing them using key statistical measures. Based on the extracted features, we then detected the 6 different causal graphs of these features with hypertension, which provide insights into the relationships and causal effects among the extracted feature variables and hypertension.
2. Significant difference analysis: Now, we need to use the corresponding 205 features to construct a causal graph under each metric. Due to the limitations of the equivalent greedy search algorithm calculation efficiency, hardware device computing power resources, and the number of subject samples, the time cost of constructing a causal graph based on 205 features is unacceptable. Therefore, we will use significant difference analysis to exclude features that do not show significant differences between hypertensive patients and healthy people. Then, considering the time cost and sample size, we will sort the retained features according to the degree of significant difference. We ultimately selected less than 50 features for causal graph construction.
3. Causal feature selection: In the following, we select the features that have a direct causal relationship with the hypertension node from the causal graph constructed under each metric. A total of 24 causal features were selected under the 6 metrics. It should be noted here that different metrics mean observing the changes of the same feature over a period of time from different perspectives. The features with the same number under different statistical metrics are essentially derivatives of the original features. Taking feature 52 as an example, we can get 4 feature variables under these metrics; they are shown in Figure 5. These 4 feature variables are essentially derivatives of feature 52. Therefore, in the final causal graph, we use feature 52 nodes to represent the above 4 features. From this, we can see that there are some features with the same number among the 24 causal features. We can use a feature node in the final causal graph to represent these feature variables with the same number, and finally obtain a final causal graph containing 10 feature nodes.

**Figure 5.** Box plot of the various statistical indicators of the feature 52.




## Strategy of Causal Inference

In order to mitigate the potential issues of bidirectional causality and cyclic graphs, we conducted the analysis of the causal relationships between respective feature variables and hypertension under each indicator, culminating in the derivation of corresponding causal subgraphs, so as to obtain the causal graph.

1. Strategy for obtaining causal graph: We randomly partitioned the dataset to identify the causal graph, with the allocation of an additional validation set for subsequent hypertension risk prediction. Recognizing that a single random partitioning could introduce undesired stochasticity (thereby rendering the resulting causal graphs potentially unrepresentative), we draw inspiration from the concept of 10-fold cross-validation. This method involves conducting 10 iterations to compute causal subgraphs, followed by a rigorous pruning process to retain only those segments demonstrating direct causal associations with hypertension within each causal subgraph. Subsequently, guided by the principle of majority rule, we amalgamate the results of these iterations to derive the ultimate causal subgraph.
2. Strategy for merging causal graph: After obtaining the final causal subgraph with each graph identified with the 6 categories of features mentioned in feature selection section, we assume that the weights of the causal relationships between the feature variables and hypertension are equal under each category of feature; based on the principle of majority rule, we integrate multiple causal subgraphs into the ultimate causal graph. This method can screen out more reliable direct causal feature variables, further simplify the causal graph, and preserve important information.

## Classifier and Performance Evaluation

In conjunction with a 10-fold cross-validation approach to partition the dataset into training and testing sets, our predictive

modeling of hypertension risk primarily leverages 4 classification algorithms: random forest, logistic regression, decision trees, and naive Bayes. These algorithms are selected for their effectiveness in capturing diverse patterns in the data. Moreover, the evaluation of our models is based on a comprehensive set of performance metrics, encompassing accuracy, precision, recall, $F_1$-score, and the AUC, which are defined later on. Following the derivation of the final causal diagram, we proceeded to select an equal number of feature variables with the strongest correlation to hypertension, based on the point-biserial correlation coefficient. These selected features were then used in the prediction of hypertension risk. Subsequently, we compared the predictive performance of this model with the one based on causal feature variables.

## Results

### Signal and Feature Analysis

We found that there are 24 feature variables directly causally related to hypertension under 6 indicators. These can be abstracted into 10 representative feature variables in the causal graph. Then, we used the point-biserial correlation coefficient to select the 24 feature variables with the strongest correlation to hypertension. After conducting data analysis, we discovered that there are 5 feature variables that overlap between the causal feature variables and the correlated feature variables. These variables are as follows and 4 of them are shown in Figure 5.

SDFeature 52 (SD of TD(sdPPGc−dPPGvalley))

QDFeature 52 (QD of TD(sdPPGc−dPPGvalley))

RFeature 52 (Range of TD(sdPPGc−dPPGvalley))

MEFeature 52 (Mean of TD(sdPPGc−dPPGvalley))

MEFeature 47 (Mean of TD(sdPPGc−PPGpeak))

Furthermore, we selected the representative samples from the groups of hypertensive patients and healthy people for comparative analysis. The PPG waveform analysis diagrams of hypertensive patients and healthy people are shown in Figure 6, and the scatter plots of feature 52 are shown in Figure 7. Then, based on the analysis of feature 52's position in PPG signals, we observed that in hypertensive patients, the peak of the c-point on sdPPG may occur earlier compared to healthy individuals. This could be a possible reason as to why feature 52 is strongly correlated with hypertension and is considered to have a strong causal relationship with hypertension.

Finally, it is important to note that further research and validation are necessary to confirm the relationship between feature 52, the c-point on sdPPG, and hypertension. These findings may provide valuable insights into potential markers for hypertension and contribute to the understanding of its pathophysiology.

**Figure 6.** Comparison of PPG waveforms between healthy people and hypertensive patients. PPG: photoplethysmogram.

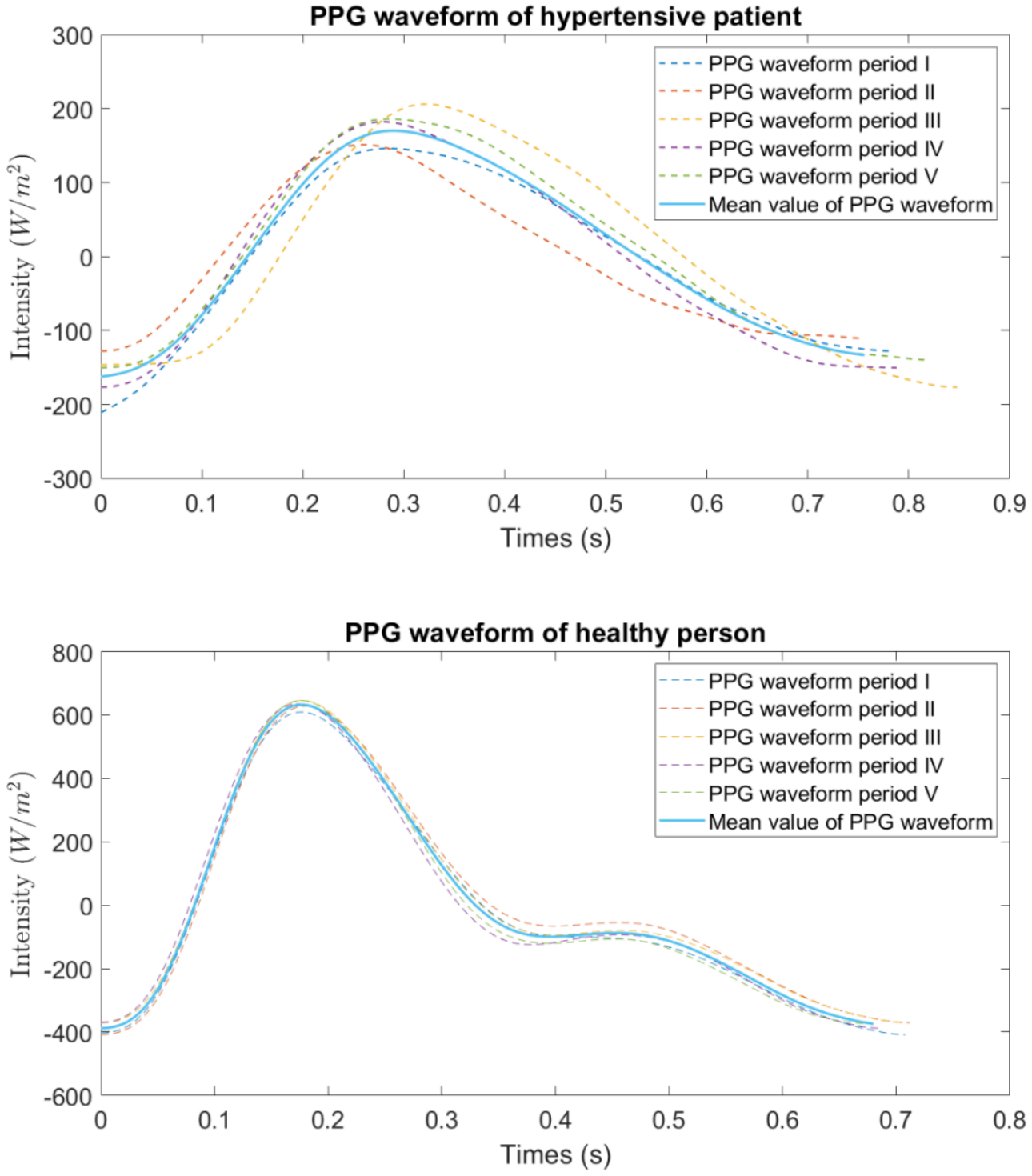
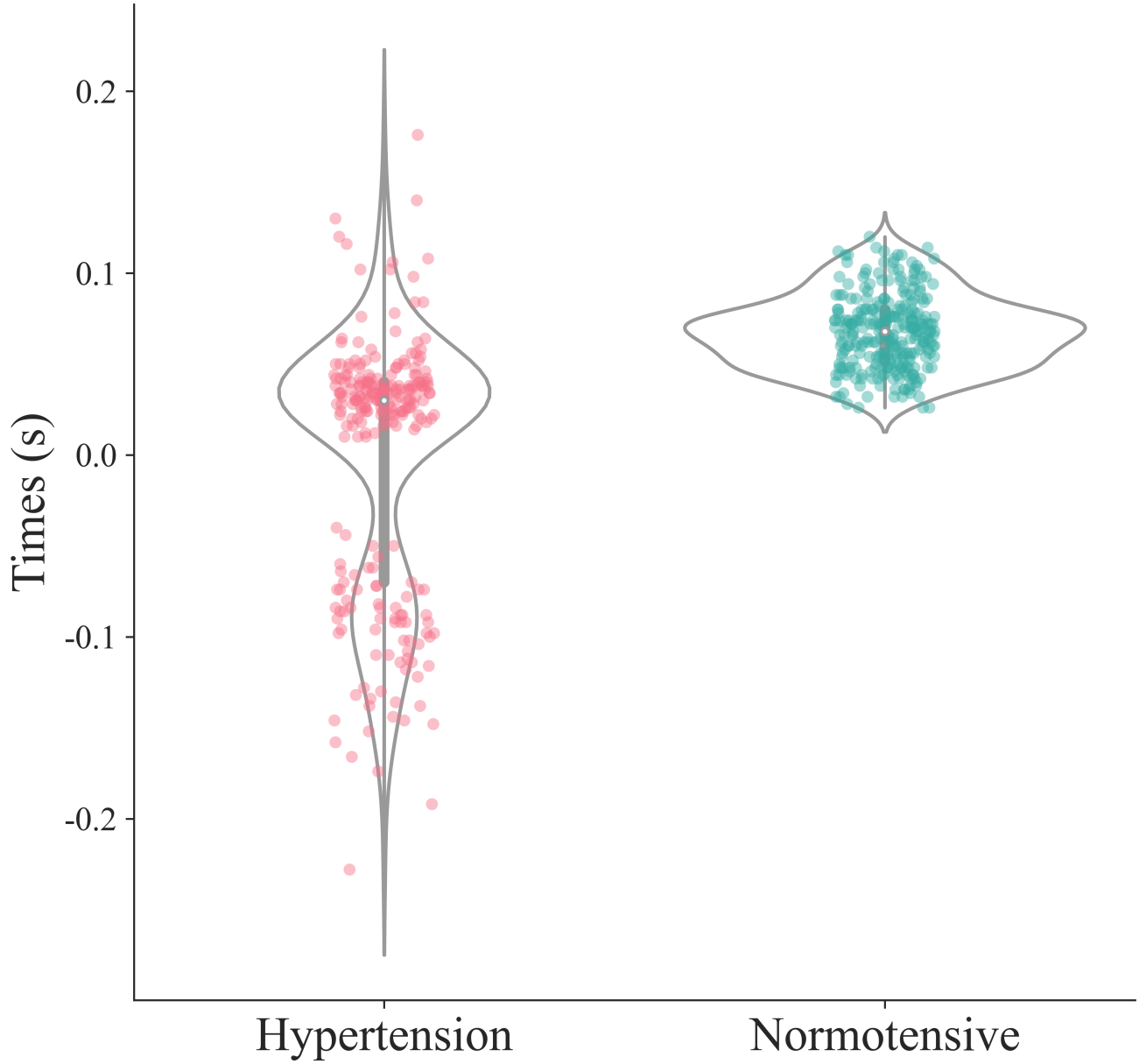


**Figure 7.** Scatter distribution of feature 52 for normotensive subjects (green) and hypertensive subjects (red).




## Causal Graph

In this study, considering the potential disturbance to the causal graph caused by randomly partitioning the data into training and testing sets, we used the idea of 10-fold cross-validation and causal strategy I to mitigate such interference. After applying the aforementioned procedures, we obtained a total of 6 causal subgraphs under different metrics. In addition, due to space constraints, this paper only presents the causal subgraphs under the standard deviation and range indicators, as shown in Figures 8 and 9, respectively. It is observed that the feature variables directly causally associated with the risk of hypertension vary across different indicators. Based on the principle of majority rule, we applied causal strategy II to obtain the final causal graph, as depicted in Figure 10.

**Figure 8.** Causal subgraph of hypertension and the features calculated with their standard deviation. AI: absolute intensity; AR: area under the PPG curve; dPPG: the first derivative of PPG; P-R: precision-recall; PPG: photoplethysmogram; RI: physiological meaningful relative index; sdPPG: the second derivative of PPG; TD: time duration.

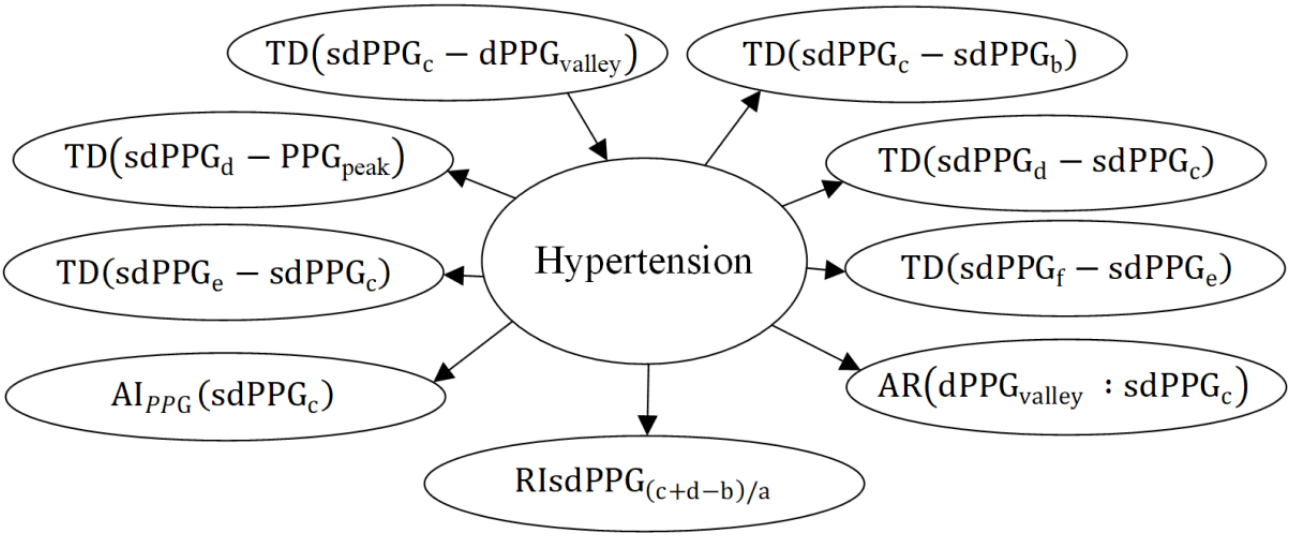



**Figure 9.** Causal subgraph of hypertension and the features calculated with their range. AM: amplitude; AR: area under the PPG curve; dPPG: the first derivative of PPG; P-R: precision-recall; PPG: photoplethysmogram; RI: physiological meaningful relative index; sdPPG: the second derivative of PPG; TD: time duration.

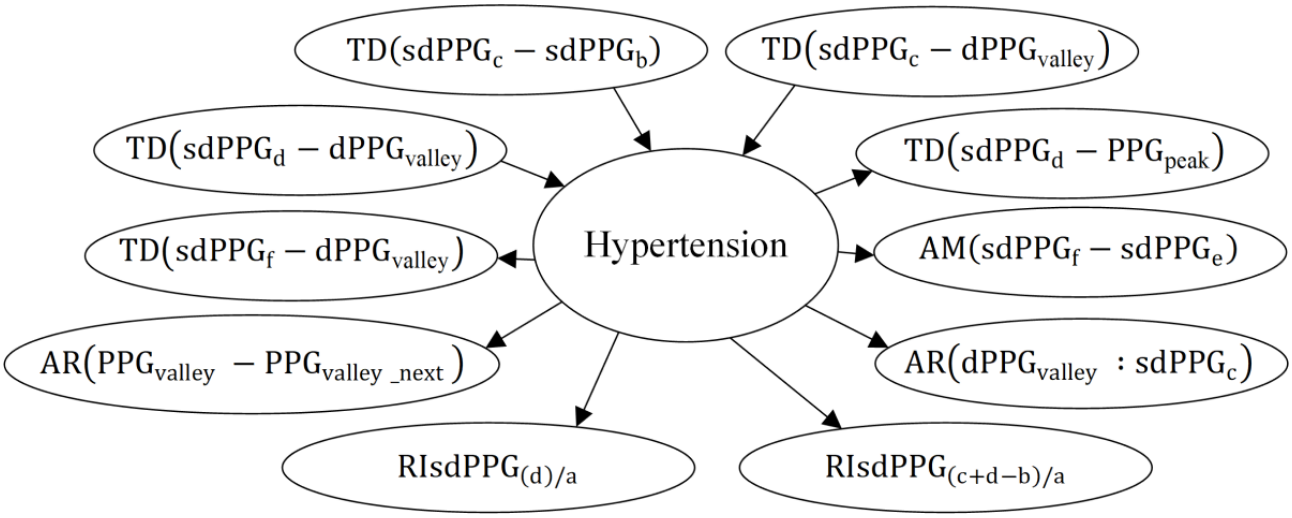
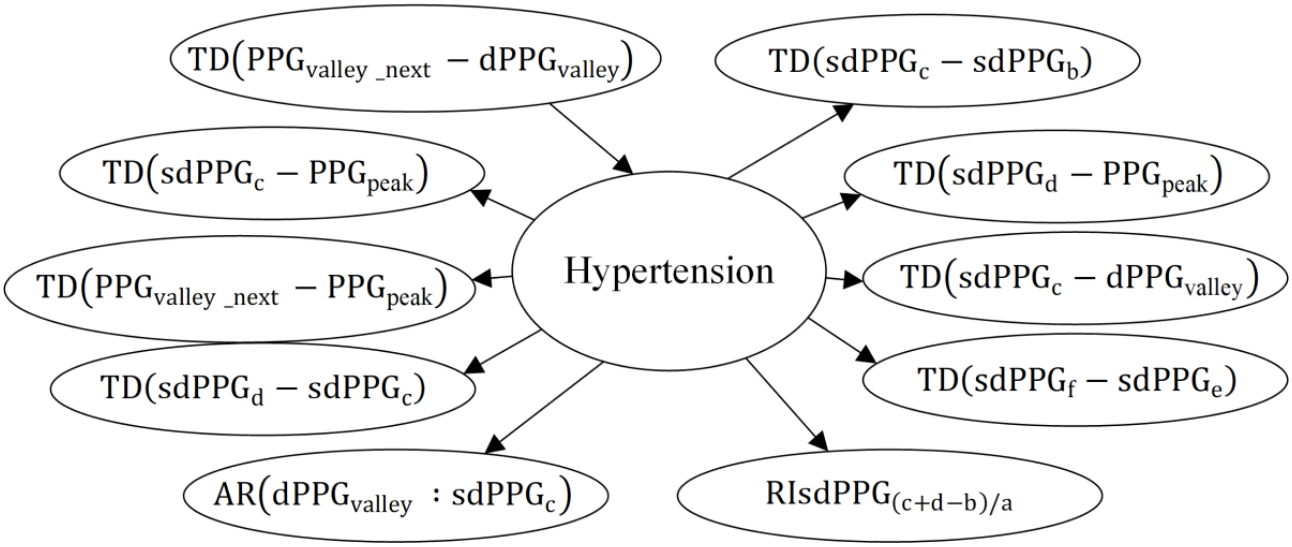



**Figure 10.** Final causal graph. AR: area under the PPG curve; dPPG: the first derivative of PPG; P-R: precision-recall; PPG: photoplethysmogram; RI: physiological meaningful relative index; sdPPG: the second derivative of PPG; TD: time duration.



### Hypertension Classification Results

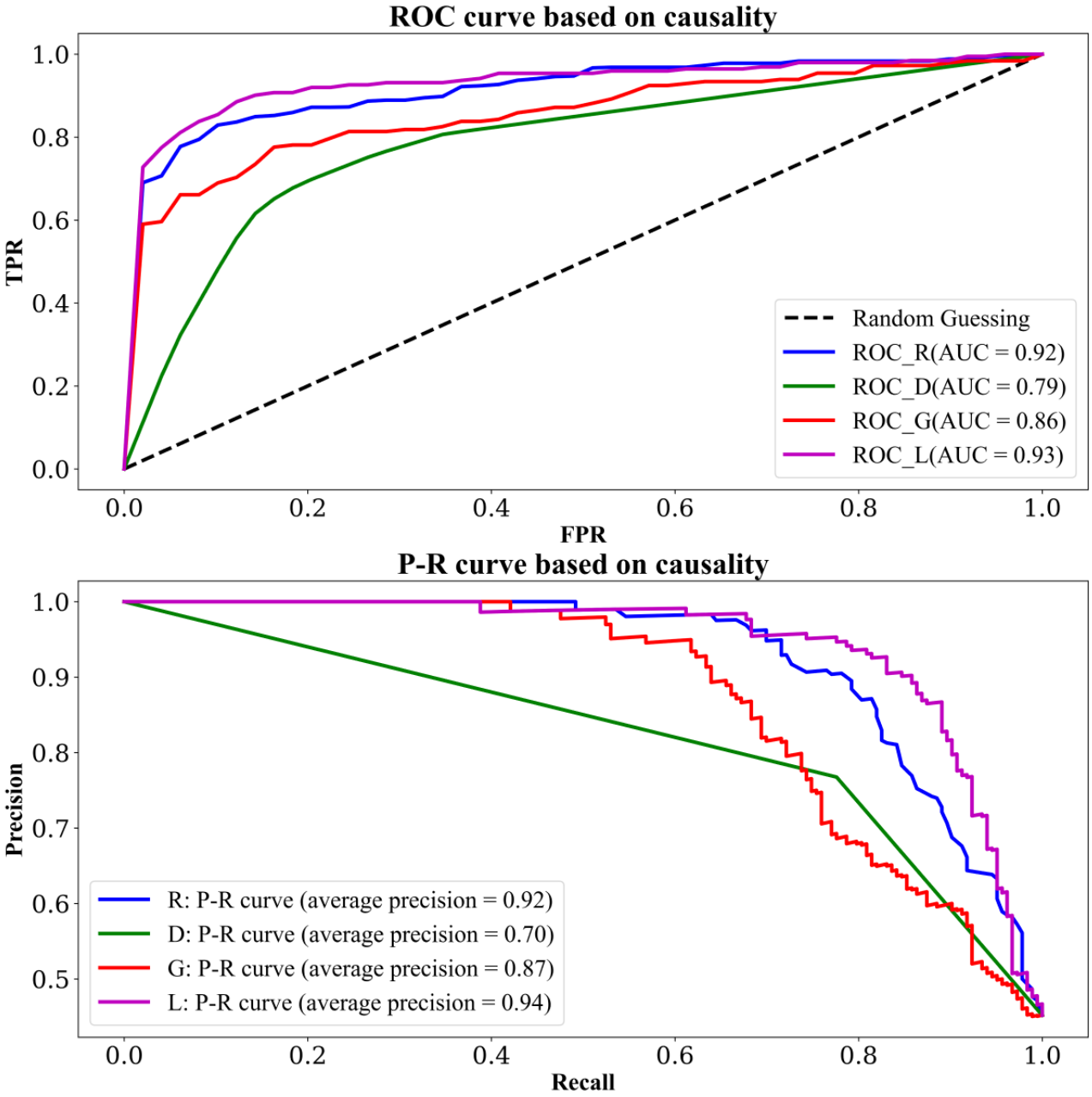
In this subsection, we used multiple classifier algorithms for hypertension classification prediction. First, we primarily utilized logistic regression and other classification algorithms based on causal feature variables for hypertension classification. The classification performance is presented in Table 2. We found that the logistic regression algorithm exhibited the best predictive performance with an accuracy of 0.89, precision of 0.92, recall of 0.82, and $F_1$-score of 0.87. Both the accuracy and accuracy rate are relatively high, which means that our classification prediction model can accurately predict hypertensive patients and healthy people, and the probability of making errors in the judgment of hypertensive patients is low; the $F_1$-score further proves the above conclusion. In addition, a higher recall rate indicates that most patients with high blood pressure can be correctly predicted.

**Table .** Causality-based classification performance.

| Algorithm | Accuracy | Precision | Recall | $F_1$-score |
|---|---|---|---|---|
| Random forest | 0.86 | 0.90 | 0.77 | 0.83 |
| Decision tree | 0.78 | 0.76 | 0.78 | 0.76 |
| Naive Bayes | 0.80 | 0.95 | 0.58 | 0.72 |
| Logistic regression | 0.89 | 0.92 | 0.82 | 0.87 |

Subsequently, Figure 11 illustrates the receiver operating characteristic curve and precision-recall curve of the classifier algorithms. The purple line represents the logistic regression classification algorithm. It can be observed that the area under the curve of this logistic regression classification algorithm is higher than that of other classification algorithms in both receiver operating characteristic and precision-recall curves.

**Figure 11.** The ROC curve (top panel) and P-R curve (bottom panel) of hypertension detection based on causal features with different machine learning algorithms: the blue curve represents random forest (R), the green curve represents decision tree (D), the red curve represents naive Bayes (G), and the purple curve represents logistic regression (L). AUC: area under the receiver operating characteristic curve; FPR: false positive rate; P-R: precision-recall; ROC: receiver operating characteristic; TPR: true positive rate.




Finally, we compared the classification performance based on causal feature variables with that based on correlated feature variables, as shown in Table 3. We found that the best performance in terms of the 4 evaluation metrics was consistently achieved by the classification algorithm based on causal feature variables. This finding is also consistent with the results presented in Figures 12 and 13. These findings imply that the causal characteristics we screened have certain mining value in the field of hypertension prediction.

**Table .** Classifier performance comparison.

| Algorithm | | Accuracy | Precision | Recall | $F_1$-score |
|---|---|---|---|---|---|
| **Causality** | | | | | |
| | Random forest | 0.86 | 0.90 | 0.77 | 0.82 |
| | Decision tree | 0.78 | 0.76 | 0.78 | 0.79 |
| | Naive Bayes | 0.80 | 0.95 | 0.58 | 0.72 |
| | Logistic regression | 0.89 | 0.92 | 0.82 | 0.87 |
| **Correlation** | | | | | |
| | Random forest | 0.79 | 0.81 | 0.72 | 0.75 |
| | Decision tree | 0.72 | 0.68 | 0.72 | 0.69 |
| | Naive Bayes | 0.80 | 0.82 | 0.74 | 0.77 |
| | Logistic regression | 0.85 | 0.88 | 0.77 | 0.82 |

**Figure 12.** The ROC curve (top panel) and P-R curve (bottom panel) for the best classifier of causality and correlation: the blue curve represents the logistic regression classifier based on causality, while the red curve represents the logistic regression classifier based on correlation. AUC: area under the receiver operating characteristic curve; FPR: false positive rate; P-R: precision-recall; ROC: receiver operating characteristic; TPR: true positive rate.

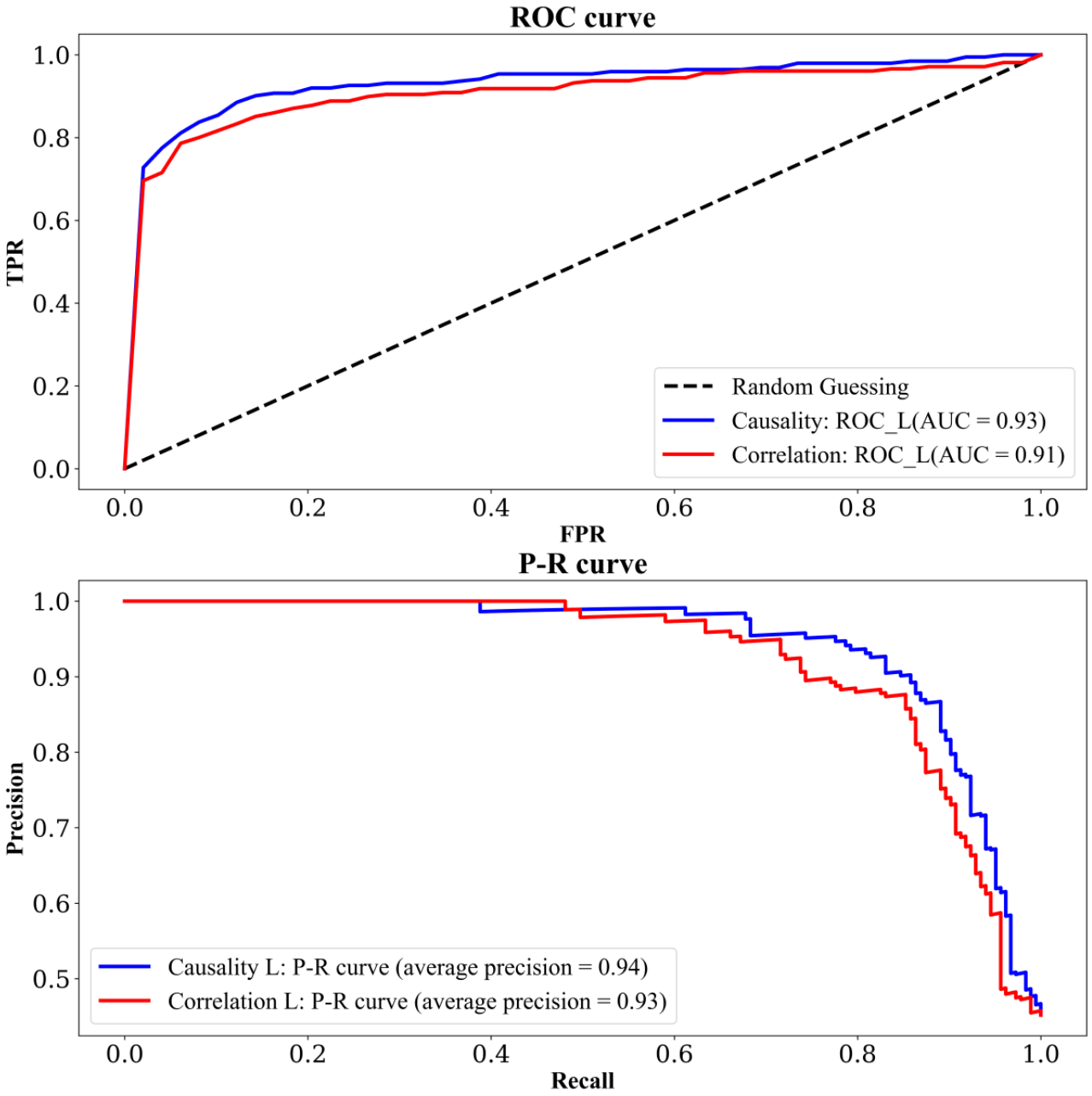
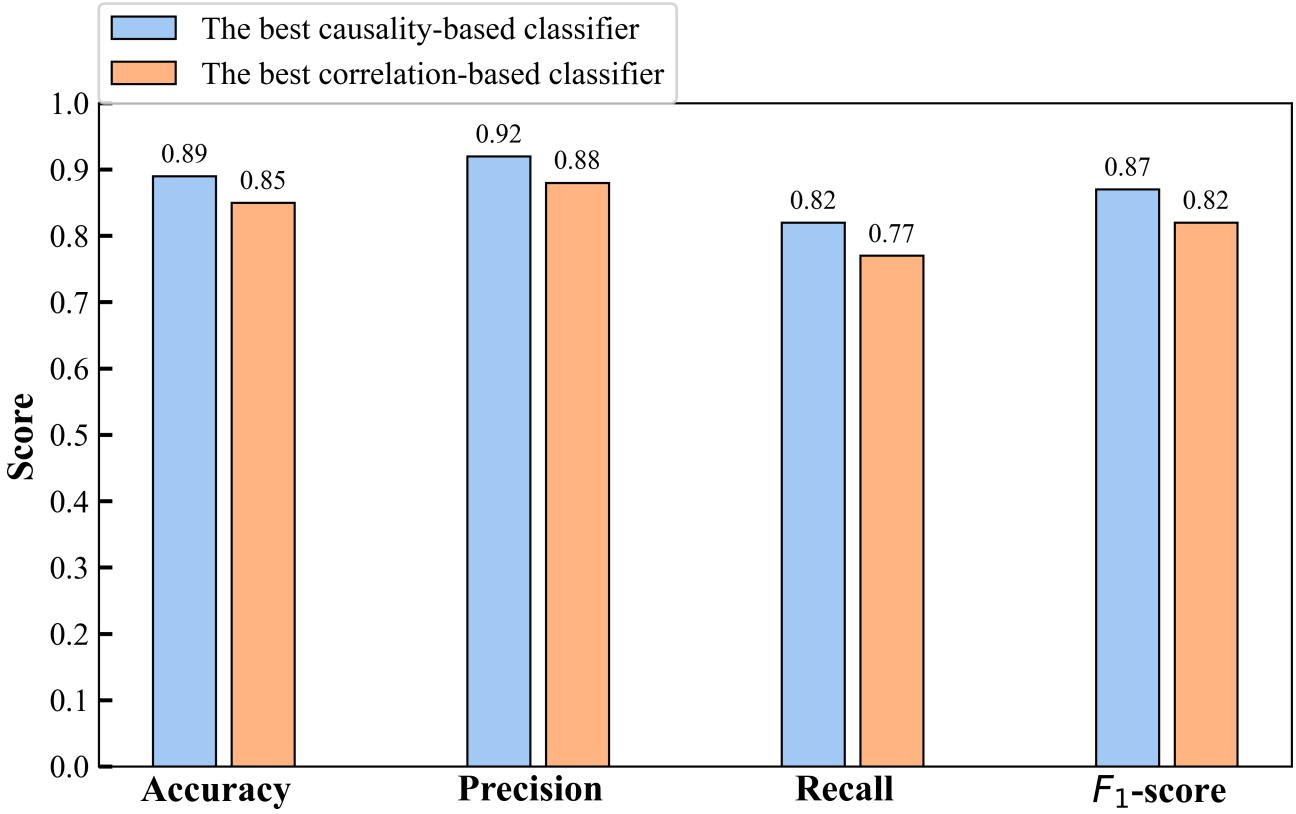


**Figure 13.** Histogram of evaluation metrics for the best classifier of causality and correlation.




## Discussion

### Principal Findings and Advantages

This study primarily explored the relationship between feature variables extracted from ECG and PPG signals and hypertension from a causal perspective, using causal inference methods to construct causal graphs. Simultaneously, to preserve the temporal information of time series signals to the maximum extent, causal graphs were constructed separately for 6 metrics, including standard deviation, mean, range, coefficient of variation, median, and quartiles. These causal graphs were derived based on specific causal strategies, ensuring a certain degree of reliability and accuracy in the resulting causal graphs. By assessing the performance of feature variables based on causality in hypertension risk classification prediction against those based on correlation, we validated the reliability of causality-based feature variables compared to correlation-based ones.

Specifically, when selecting feature variables strongly associated with hypertension, both causal inference and correlation coefficient–based methods performed similarly. However, when the association between feature variables and hypertension was weak, causal inference methods tended to select more reliable feature variables compared to correlation-based methods. This is the reason why feature variables based on causality outperformed those based on correlation in hypertension risk prediction. Additionally, we found that feature 52's derived variables exhibited significant differences in distribution between the hypertensive and healthy subject groups under multiple metrics. This may provide potential value and insights for subsequent pathological mechanism analysis.

### Comparison to Prior Work

This study conducted exploratory analysis, initially focusing on the correlation analysis between hypertension and blood pressure based on the medical information mart for intensive care (MIMIC) database. Typically, the gold standard for diagnosing hypertension is SBP and DBP, where subjects are considered hypertensive when SBP exceeds 140 mm Hg or DBP exceeds 90 mm Hg. Nevertheless, when clustering analysis was performed on 24-hour dynamic blood pressure data collected from patients, we observed that the blood pressure distribution of hypertensive and nonhypertensive subjects did not exhibit significant differentiation or stratification; instead, they appeared mixed. After analysis, we attributed this phenomenon to factors such as patients taking antihypertensive medications, being in specific states, or incorrect device wear, which indirectly reflects the limitations of blood pressure measurement. Second, we previously conducted causal analysis [16] using data collected from a self-generated database of 30 individuals. Causal analysis was primarily carried out under the mean metric, resulting in limited preservation of temporal information. However, it still revealed significant differences in the distribution of feature 52 between the hypertensive and healthy subject groups, consistent with the findings of this paper.

### Limitations and Future Work

There were some limitations to this study. First, our work primarily focused on binary classification to distinguish hypertensive patients from healthy individuals. However, hypertension can be categorized into different stages, such as stage 1, stage 2, and stage 3, based on blood pressure level and disease condition. Second, the population used could have been more diverse in terms of race and ethnicity. In our future work,

we will consider conducting clustering of the features to distinguish different stages of hypertension, and we will validate the work on larger and more diverse subject populations to be able to draw more general conclusions.

## Conclusion

In this study, we explored the feasibility of predicting the risk of hypertension using causal inference methods. First, we constructed causal graphs using the GES algorithm and 10-fold cross-validation approach under each indicator. We then applied corresponding causal strategies to obtain the optimal causal graphs for each indicator. Finally, we merged the causal graphs from different indicators into a final causal graph based on the majority rule. After selecting the feature variables, we used classifiers including random forests, decision trees, naive Bayes, and logistic regression to predict hypertension. Overall, combining various indicators, we found that most classifiers based on causal features have better classification performance than classifiers based on correlation features. To the best of our knowledge, this study represents the first attempt to introduce causal inference methods in hypertension prediction, providing a new perspective for understanding the physiological mechanisms of hypertension.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (82102178) and Huzhou ST Special Program of Huzhou (2023GZ01).

## Data Availability

The data we used for this study came from a public dataset, which other researchers can apply to access [17].

## Authors' Contributions

KG's contributions include data curation, formal analysis, investigation, methodology, software, validation, visualization, and writing the original draft. YC's contributions include conceptualization, resources, and reviewing and editing the manuscript. XS and ZF contributed to visualization. XD's contributions include conceptualization, funding acquisition, methodology, project administration, resources, supervision, and reviewing and editing the manuscript.

## Conflicts of Interest

None declared.

## References

1. Williams B, Mancia G, Spiering W, et al. ESC/ESH Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). Eur Heart J 2018;39(33):3021-3104. [doi: 10.1093/eurheartj/ehy339]
2. Jain P, Gajbhiye P, Tripathy RK, Acharya UR. A two-stage deep CNN architecture for the classification of low-risk and high-risk hypertension classes using multi-lead ECG signals. Inform Med Unlocked 2020;21:100479. [doi: 10.1016/j.imu.2020.100479]
3. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. Nat Rev Nephrol 2020 Apr;16(4):223-237. [doi: 10.1038/s41581-019-0244-2] [Medline: 32024986]
4. Mancia G, Kreutz R, Brunström M, et al. ESH Guidelines for the management of arterial hypertension. The Task Force for the management of arterial hypertension of the European Society of Hypertension: endorsed by the International Society of Hypertension (ISH) and the European Renal Association (ERA). J Hypertens 2023 Dec 1;41(12):1874-2071. [doi: 10.1097/HJH.0000000000003480] [Medline: 37345492]
5. Huang QF, Yang WY, Asayama K, et al. Ambulatory blood pressure monitoring to diagnose and manage hypertension. Hypertension 2021 Feb;77(2):254-264. [doi: 10.1161/HYPERTENSIONAHA.120.14591] [Medline: 33390042]
6. Paragliola G, Coronato A. An hybrid ECG-based deep network for the early identification of high-risk to major cardiovascular events for hypertension patients. J Biomed Inform 2021 Jan;113:103648. [doi: 10.1016/j.jbi.2020.103648] [Medline: 33276113]
7. Elgendi M, Fletcher R, Liang Y, et al. The use of photoplethysmography for assessing hypertension. NPJ Digit Med 2019;2:60. [doi: 10.1038/s41746-019-0136-7] [Medline: 31388564]
8. Alkhodari M, Islayem DK, Alskafi FA, Khandoker AH. Predicting hypertensive patients with higher risk of developing vascular events using heart rate variability and machine learning. IEEE Access 2020;8:192727-192739. [doi: 10.1109/ACCESS.2020.3033004]
9. Rajput JS, Sharma M, Tan RS, Acharya UR. Automated detection of severity of hypertension ECG signals using an optimal bi-orthogonal wavelet filter bank. Comput Biol Med 2020 Aug;123:103924. [doi: 10.1016/j.compbiomed.2020.103924] [Medline: 32768053]

10. Pearl J. Comment: Understanding Simpson's Paradox. In: Probabilistic and Causal Inference: The Works of Judea Pearl 2022:399-412. [doi: 10.1145/3501714.3501738]
11. Yao L, Chu Z, Li S, Li Y, Gao J, Zhang A. A survey on causal inference. ACM Trans Knowl Discov Data 2021 Oct 31;15(5):1-46. [doi: 10.1145/3444944]
12. Mieloszyk R, Twede H, Lester J, et al. A comparison of wearable tonometry, photoplethysmography, and electrocardiography for cuffless measurement of blood pressure in an ambulatory setting. IEEE J Biomed Health Inform 2022 Jul;26(7):2864-2875. [doi: 10.1109/JBHI.2022.3153259] [Medline: 35201992]
13. Ding X, Yan BP, Zhang YT, et al. Feature exploration for knowledge-guided and data-driven approach based cuffless blood pressure measurement. arXiv. Preprint posted online on Aug 27, 2019 URL: https://arxiv.org/abs/1908.10245 [accessed 2024-12-03]
14. Chickering DM. Optimal structure identification with greedy search. J Mach Learn Res 2002;3(Nov):507554. [doi: 10.1162/153244303321897717]
15. Huang B, Zhang K, Lin Y, Schölkopf B, Glymour C. Generalized score functions for causal discovery. KDD 2018 Aug;2018:1551-1560. [doi: 10.1145/3219819.3220104] [Medline: 30191079]
16. Gong K, Chen Y, Ding X. Causal inference for hypertension prediction. 2023 Presented at: 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); Jul 24-27, 2023; Sydney, Australia. [doi: 10.1109/EMBC40787.2023.10341021]
17. Aurora. Microsoft. URL: https://www.microsoft.com/en-us/research/project/aurora/ [accessed 2025-01-07]

## Abbreviations

**AI:** absolute intensity
**AM:** amplitude
**AR:** area under the PPG curve
**AUC:** area under the receiver operating characteristic curve
**BP:** blood pressure
**DAG:** direct acyclic graph
**DBP:** diastolic blood pressure
**dPPG:** the first derivative of PPG
**ECG:** electrocardiogram
**GES:** greedy equivalence search
**P-R:** precision-recall
**PPG:** photoplethysmogram
**PTT:** pulse transit time
**RI:** physiological meaningful relative index
**RRI:** R-R interval
**SBP:** systolic blood pressure
**sdPPG:** the second derivative of PPG
**TD:** time duration

*Edited by A Coristine; submitted 06.05.24; peer-reviewed by A Jain, L Vlad, X Tian, X Xiao; revised version received 21.10.24; accepted 21.10.24; published 23.01.25.*

*Please cite as:*
*Gong K, Chen Y, Song X, Fu Z, Ding X*
*Causal Inference for Hypertension Prediction With Wearable Electrocardiogram and Photoplethysmogram Signals: Feasibility Study*
*JMIR Cardio 2025;9:e60238*
*URL: https://cardio.jmir.org/2025/1/e60238*
*doi:10.2196/60238*

Original Paper

# Efficacy of Unsupervised YouTube Dance Exercise for Patients With Hypertension: Randomized Controlled Trial

Mizuki Sakairi[1], MD; Taiju Miyagami[1], PhD; Hiroki Tabata[2], PhD; Naotake Yanagisawa[3], PhD; Mizue Saita[1], PhD; Mai Suzuki[1], PhD; Kazutoshi Fujibayashi[1], PhD; Hiroshi Fukuda[1], PhD; Toshio Naito[1], PhD

[1]Department of General Medicine, Faculty of Medicine, Juntendo University, Tokyo, Japan

[2]Juntendo Advanced Research Institute for Health Science, Tokyo, Japan

[3]Medical Technology Innovation Center, Juntendo University, Tokyo, Japan

**Corresponding Author:**
Taiju Miyagami, PhD
Department of General Medicine
Faculty of Medicine
Juntendo University
2-1-1 Hongo
Bunkyo-ku
Tokyo, 113-0033
Japan
Phone: 81 338133111
Email: tmiyaga@juntendo.ac.jp

## Abstract

**Background:** High blood pressure (BP) is linked to unhealthy lifestyles, and its treatment includes medications and exercise therapy. Many previous studies have evaluated the effects of exercise on BP improvement; however, exercise requires securing a location, time, and staff, which can be challenging in clinical settings. The antihypertensive effects of dance exercise for patients with hypertension have already been verified, and it has been found that adherence and dropout rates are better compared to other forms of exercise. If the burden of providing dance instruction is reduced, dance exercise will become a highly useful intervention for hypertension treatment.

**Objective:** This study aims to investigate the effects of regular exercise therapy using dance videos on the BP of patients with hypertension, with the goal of providing a reference for prescribing exercise therapy that is highly feasible in clinical settings.

**Methods:** This nonblind, double-arm, randomized controlled trial was conducted at Juntendo University, Tokyo, from April to December 2023. A total of 40 patients with hypertension were randomly assigned to either an intervention group (dance) or a control group (self-selected exercise), with each group comprising 20 participants. The intervention group performed daily dance exercises using street dance videos (10 min per video) uploaded to YouTube. The control group was instructed to choose any exercise other than dance and perform it for 10 minutes each day. The activity levels of the participants were monitored using a triaxial accelerometer. BP and body composition were measured on the day of participation and after 2 months. During the intervention period, we did not provide exercise instruction or supervise participants' activities.

**Results:** A total of 34 patients were included in the study (16 in the intervention group and 18 in the control group). The exclusion criteria were the absence of BP data, medication changes, or withdrawal from the study. The mean age was 56 (SD 9.8) years, and 18 (53%) of the patients were female. The mean BMI was 28.0 (SD 6.3) m/kg$^2$, and systolic blood pressure (SBP) and diastolic blood pressure (DBP) were 139.5 (SD 17.1) mm Hg and 85.8 (SD 9.1) mm Hg, respectively. The basic characteristics did not differ between the two groups. In the multivariate analysis, SBP and DBP improved significantly in the intervention group compared to the control group (mean SBP –12.8, SD 6.1 mm Hg; *P*=.047; mean DBP –9.7, SD 3.3 mm Hg; *P*=.006).

**Conclusions:** This study evaluated the effects of dance exercise on patients with hypertension, as previously verified, under the additional condition of using dance videos without direct staff instruction or supervision. The results showed that dance videos were more effective in lowering BP than conventional exercise prescriptions.

**Trial Registration:** University Hospital Medical Information Network UMIN 000051251; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000058446

*(JMIR Cardio 2025;9:e65981)* doi:10.2196/65981

**KEYWORDS**

dance; video; exercise therapy; hypertension; blood pressure therapy; YouTube; mHealth

## Introduction

High blood pressure (BP) is a major chronic disease that threatens people's health and is an important risk factor for many types of heart, brain, and kidney vascular diseases. A total of 590,000 Japanese individuals with high BP continuously receive medical care, the highest number among lifestyle-related diseases [1]. The prevalence of high BP among adults in the United States was 29% from 2011 to 2014, and the prevalence rates increased with age: 18-39 years, 7.3%; 40-59 years, 32.2%; and 60 years and older, 64.9% [2]. The global population aged older than 65 years is expected to double between 2019 and 2050 [3]. Japan has the oldest population worldwide; in 2013, those aged older than 65 years exceeded 25% of the population and are expected to exceed 40% by 2060 [4]. Therefore, high BP is a global public health problem, and the number of patients with the condition is expected to increase with the growth of the aging population.

High BP is associated with an unhealthy lifestyle. The clinical treatment of high BP involves antihypertensive medications and lifestyle interventions, such as reducing salt intake, eating a diet rich in fruits and vegetables, exercising, and maintaining a healthy body weight [5]. Although antihypertensive medications are the main treatment, exercise is also an important recommendation for patients with high BP [6-8]. It is known that regular moderate exercises, such as water walking, brisk walking, running, small-sided soccer, and swimming, have beneficial effects on BP in patients with hypertension [9-13]. The World Health Organization recommends at least 150 minutes of moderate to vigorous physical activity (MVPA) per week [14]. However, in Japan, only about half of the population (59.6% of men and 46.9% of women) meets these physical activity standards [15]. Furthermore, during the COVID-19 pandemic, restrictions on outdoor activities led to decreased physical activity levels [16]. It has also been suggested that safety concerns, especially for women when exercising alone outdoors or after sunset, as well as fear of criticism, are barriers to engaging in physical activity [17]. Challenges in securing time and space for exercise due to caregiving, childcare, employment, and pandemics hinder physical activity. Furthermore, although physical activity interventions delivered or prompted by health professionals in primary care appear effective in increasing participation in MVPA, exercise prescription training for health care professionals is inadequate [18].

Dance, a fun form of exercise that uses music and can be performed in confined spaces, remains feasible, even in situations such as the COVID-19 pandemic. Dance was part of Japan's educational curriculum in 2012 and was added as an Olympic sport starting in 2024 [19]. A survey conducted in Japan indicated that the proportion of teenagers participating in hip-hop dance at least once a week rose from 2.1% in 2015 to 3.5% in 2023 [20]. Therefore, dance has become an accessible sport, and compared to other activities such as marathon running or swimming, is easier for patients to perform in terms of space and time. A meta-analysis comparing dance to other exercises found that adherence and dropout rates for dance were better than those for other forms of exercise [21]. Previous studies have shown that regular dance therapy can benefit hypertension management in patients [22-30]. However, to the best of our knowledge, no studies in Japan have examined the effects of dance on BP. Additionally, previous studies involved direct patient monitoring during exercise or used internet-based methods for monitoring. In clinical settings, it is challenging to gather participants for regular prescribed group dance sessions or to monitor them using video chat. We, therefore, aimed to investigate the effect of regular dance therapy interventions on BP in patients with hypertension to provide a reference for prescription studies on dance exercise therapy in these patients. We hypothesized that performing the same movements without monitoring using self-made dance videos could lower BP and be useful as a nonpharmacological treatment for high BP.

## Methods

### Ethical Considerations

This study was approved by the Ethics Committee of Juntendo University (approval: E22-0387). The participants received written information about the trial, including its aim, expected advantages, and role, and were asked to provide written informed consent. This study was retrospectively registered with the University Hospital Medical Information Network (UMIN) under ID UMIN 000051251 and with the International Standard Randomized Controlled Trial Number registry (under ID ISRCTN46013). The UMIN is a network member of the Japan Primary Registries Network, as described in the World Health Organization registry network. All procedures were performed in accordance with relevant guidelines and regulations.

### Setting and Design

This study was conducted at the Juntendo University Department of General Medicine, Tokyo, Japan, a regional core hospital that treats many patients with lifestyle-related diseases. Outpatients generally visit the hospital every 2 months.

This was a nonblind, double-arm randomized controlled trial conducted from April 1, 2023, to December 27, 2023. Based on a previous study [31], we set the intergroup difference (difference from baseline) to –9 and the SD at 9. The results of previous studies are as follows: mean difference (MD) –8.75 mm Hg; 95% CI –6.51 to –10.39 for systolic BP, and MD –8.35 mm Hg; 95% CI –6.25 to –10.45 for diastolic BP. This study anticipated a similar decrease in BP, as reported previously. With a desired power of 80%, a sample size of 34 individuals was calculated. Considering a dropout rate of 15%, we selected a sample size of 40 participants, allocated in a 1:1 ratio into two groups using a random number table: the intervention (dance)

group (n=20) and the control group (n=20). TM created the randomization table, staff members (MSakairi) conducted the recruitment, and the admin assistant conducted the group allocation.

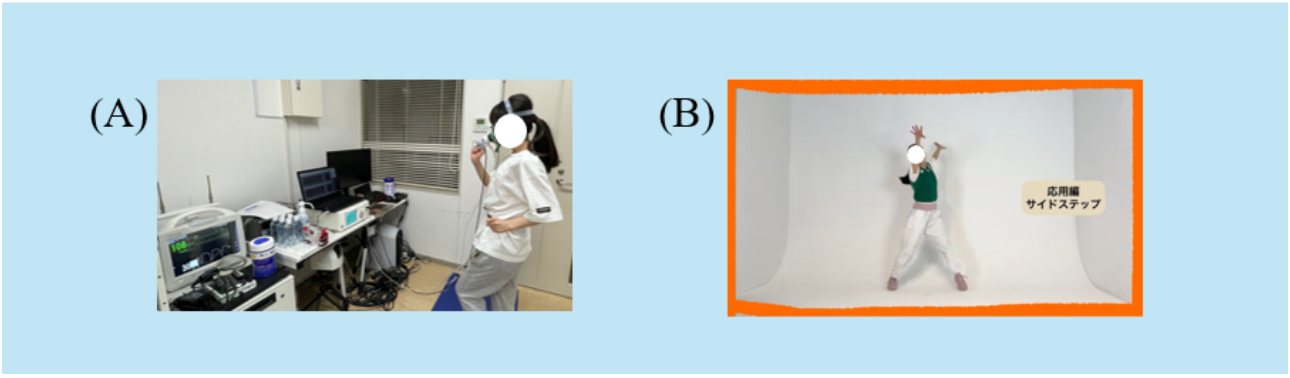
We included outpatients with high BP from the Juntendo University Department of General Medicine. These patients with hypertension had been diagnosed with hypertension and were receiving regular oral medication. The patient was invited to participate in this study by their primary physician, whom they regularly visited for hypertension management, and consent was obtained. Participants were informed that their participation in this study was voluntary and that they could withdraw if they chose to discontinue after joining. Additionally, if their primary physician determined that withdrawal was necessary due to changes in their medical condition, the study could be terminated. We excluded patients with complications rendering them unsuitable for exercise, such as cardiovascular disease, cerebral vascular disease, those unable to balance on one leg, and patients who were newly prescribed antihypertensive drugs or who were administered antihypertensives later.

## Interventions

### *Development of Dance Videos*

The intervention group watched an approximately 10-minute-long dance video and replicated the movements. The dance videos for the intervention group were created using the following materials and procedures. One of the authors (MSakairi), with 29 years of extensive experience in dance, developed a dance program based on street dance, with reference to instructional videos for school classes [32]. The music used for the dance was selected from DOVA-SYNDROME [33]. The staff used exhaled-breath analysis to measure the dance activity level and create five videos ranging from 4.5 to 7 metabolic equivalent of task (METs), measuring the intensity of physical activity that represents the metabolic rate relative to the resting metabolic rate (Figure 1). The formula used to calculate METs is expressed as follows:

**Figure 1.** Details about dance. (A) The process of creating the dance. We have used exhaled breath analysis to measure the activity level of dance and created five videos ranging from 4.5 to 7 METs. (B) A part of the distributed dance video. We distributed the video of the dance we created to participants using YouTube. MET: metabolic equivalent of task.




During the dance activity, METs were measured using a respiratory gas analyzer (pulmonary exercise load monitoring system: AE-310S, Minato Medical Science Co, Ltd, Osaka city, Osaka, Japan). The average METs for each dance video were as follows: (1) 4.57, (2) 4.86, (3) 4.84, (4) 6.95, and (5) 7.11 METs. Measurements were conducted using the breath-by-breath method to calculate $VO_2$ and $VCO_2$ based on signals from high-precision flow sensors [34]. We uploaded the created dance videos to YouTube with restricted access.

### *Intervention Group Procedures*

On the day of recruitment, we provided the intervention group with a URL to access the five YouTube videos. Participants were instructed to freely select a dance from the 5 videos and perform it daily while watching the video. We did not provide any guidance on dance instruction or supervision during the dance sessions. However, we instructed the control group to freely select any exercise other than dance and perform it for 10 minutes daily. Additionally, on the day of recruitment, BP and body composition were measured, and web-based surveys were administered using Google Forms to all participants. BP was measured using an automatic medical electronic BP monitor (HBP-9035 Kentaro, OMRON Health Care Co, Ltd, Kyoto City, Kyoto Prefecture, Japan).

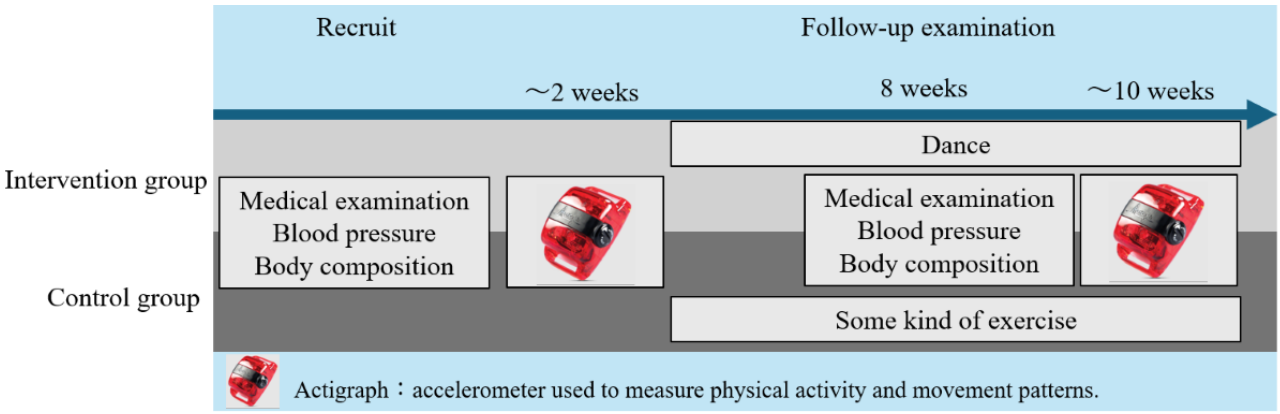
Participants from both groups were instructed not to change their lifestyle 2 weeks from the day of recruitment and to wear an ActiGraph continuously during this period, except during sleep and bathing. ActiGraph is a 3-axis accelerometer (wGT3X-BT ActiGraph, ActiGraph, LLC). Actigraph triaxial accelerometers are the most extensively used devices in numerous studies focused on monitoring human physical activity energy expenditure; they are capable of detecting changes in motion and converting them into digital signals, which can then be analyzed to estimate energy expenditure [35].

Two weeks after recruitment, both the intervention and control groups were instructed to begin their designated exercises and continue until the end of the study period.

Approximately 2 months after recruitment, during a regular outpatient visit, BP and body composition were measured again, and another web-based survey was completed. Subsequently, the participants were instructed to wear the ActiGraph continuously, except during sleep and bathing, for another 2 weeks (Figure 2). During the intervention period, participants in both the intervention and control groups did not receive

exercise guidance, nor were the frequency or manner of their exercise monitored. We did not compensate the participants of this study. The research data of patients in this study were anonymized using identification numbers; however, researchers could still identify individual patients with these numbers.

**Figure 2.** Research schedule. We instructed both the intervention group and the control group to exercise and measured their physical activity levels using an actigraph.




## Outcome Measures

### *Variables*

The variables used in this study were gender, age, number of antihypertensive drugs, number of lifestyle-related diseases (diabetes, dyslipidemia, and hyperuricemia), medical history (cerebral infarction and ischemic heart disease), height, body weight, body muscle mass, body fat mass, family in need of care (children and adults), the presence of cohabitants, exercise habits, systolic blood pressure (SBP), diastolic blood pressure (DBP), and MVPA per day (corresponding to activity levels that are moderate or higher in intensity, namely, a level of 3 METs or higher).

### *Primary Outcome*

The main outcome of this study was BP. During the study period, we measured the BP and body composition of the patients twice for comparison. This was performed on the day of participation and 2 months after participation during outpatient visits.

## Data Collection

We obtained the participants' gender, age, frequency of antihypertensive medication use, lifestyle-related diseases (diabetes, dyslipidemia, and hyperuricemia), and medical history (cerebral infarction, and ischemic heart disease) from medical records for both groups. The body composition measured on the day of recruitment and 2 months later included height, weight, muscle mass, and body fat mass. In addition, a web-based survey using Google Forms was conducted to inquire about the presence of cohabitants, caregivers (both children and adults), and exercise habits. The criteria of the ActiGraph for adopting the data involved confirming valid days with worn durations of 10 hours or more per day, with at least 7 such days within 2 weeks. The average value for the adopted days was calculated for each individual [36-38]. In this study, as it is exploratory research rather than a confirmatory study, we did not perform multiplicity adjustments.

## Statistical Analyses

All statistical analyses were performed using JMP Pro (version 16.0; SAS Institute). All reported *P* values were 2-tailed, and *P* values <.05 were considered statistically significant. The results are presented as mean (SD) for continuous variables or as prevalence (%) for categorical variables. Comparisons between two groups were performed using the chi-square test. Multiple regression analysis was performed on both groups, with BP as the dependent variable. The other covariates were gender, age, and daily MVPA before starting exercise.

# Results

A total of 40 patients participated in the study (see Multimedia Appendix 1 for CONSORT [Consolidated Standards of Reporting Trials] checklist), and 20 outpatients were evaluated in each intervention and control group. We excluded 2 patients who lacked BP data, one patient who changed medications, and 1 patient who withdrew to care for a parent from the dance group. We also excluded one patient who changed medications and one patient who took a double dose from the control group. These participants could have experienced BP changes due to antihypertensive medications, and the lack of BP data makes evaluation difficult. Including these participants may reduce validity, so it is reasonable to exclude them. Therefore, 16 patients in the intervention group and 18 patients in the control group were analyzed (Figure 3). Among the participants, 18 (53%) participants were female, 4 (12%) participants were family caregivers, and 19 (56%) participants had lifestyle diseases (diabetes, dyslipidemia, and hyperuricemia). The mean age was 56 (SD 9.8) years, the mean number of patients who took treatment with an antihypertensive drug was 1.5 (SD 0.5), the mean BMI was 28.0 (6.3) m/kg$^2$, the mean body muscle mass was 46.5 (SD 9.6) kg, the mean body fat mass was 25.3 (SD 13.8) kg, the mean MVPA time of per day was 20.8 (SD 14.3) minutes, and the mean SBP and DBP were 139.5 (SD 17.1) and 85.8 (SD 9.1) mm Hg (Table 1).

**Figure 3.** Number of participants and exclusions from the study. Four participants were excluded from the intervention group and two from the control group.

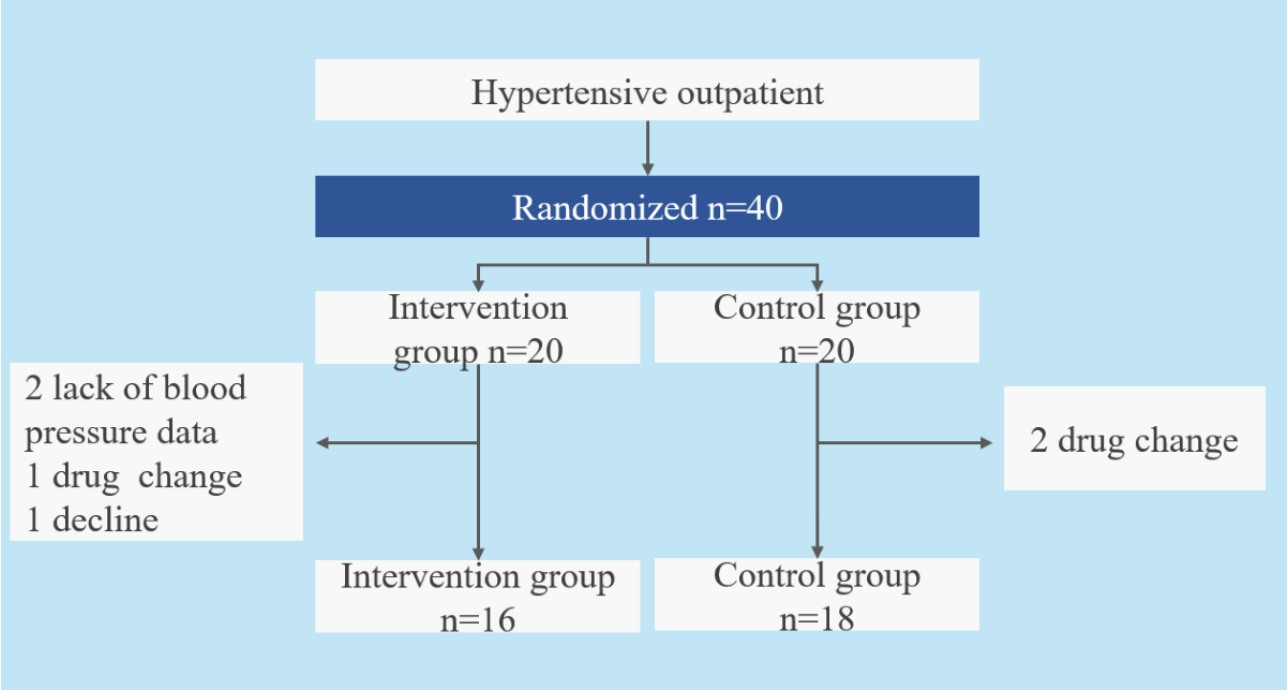



**Table 1.** Characteristics comparing intervention and control groups[a].

| Variable | Total (n=34) | Intervention group (n=16) | Control group (n=18) | *P* value |
|---|---|---|---|---|
| Sex (female), n (%) | 18 (53) | 9 (56) | 9 (50) | .70 |
| Age (years), mean (SD) | 56 (9.8) | 54 (11) | 59 (8) | .20 |
| Antihypertensive drug, mean (SD) | 1.5 (0.5) | 1.5 (0.5) | 1.5 (0.5) | .10 |
| Lifestyle disease, n (%) | 19 (56) | 8 (50) | 11 (61) | .50 |
| BMI (m/kg$^2$), mean (SD) | 28.0 (6.3) | 27.2 (1.5) | 29.1 (1.5) | .80 |
| Body muscle mass (kg), mean (SD) | 46.5 (9.6) | 45.4 (9.7) | 47.4 (9.7) | .60 |
| Body fat mass (kg), mean (SD) | 25.3 (13.8) | 23.0 (14.0) | 27.1 (13.9) | .40 |
| Family caregiver, n (%) | 4 (12) | 2 (13) | 2 (11) | .90 |
| SBP[b] (mm Hg), mean (SD) | 139.5 (17.1) | 141 (4.6) | 138.2 (3.8) | .60 |
| [c]DBP (mm Hg), mean (SD) | 85.8 (9.1) | 86.3 (11.4) | 85.4 (6.8) | .80 |
| MVPA[d] per day (minutes), mean (SD) | 20.8 (14.3) | 24.7 (4.6) | 17.3 (9.7) | .20 |

[a]This is the blood pressure measured on the first day of recruitment.

[b]SBP: systolic blood pressure.

[c]DBP: diastolic blood pressure.

[d]MVPA: moderate to vigorous physical activity (moderate intensity activities range from 3.0 to 5.9 METs, while high-intensity activities are 6.0 METs or above).

As a result, there was a difference in SBP between the 2 groups. The mean for the intervention group was –7.9 (SD 18.1) mm Hg and the mean for the control group was 3.9 (SD 14.5) mm Hg (*P*=.04). No difference was observed in DBP (mean –6.6, SD 11.1 mm Hg; mean –0.94, SD 10.6 mm Hg; *P*=.14), body weight (mean –3.5, SD 13.3 kg; mean –5.4, SD 18.7 kg; *P*=.74), body muscle mass (mean –7.9, SD 16.6 kg; mean –5.1, SD 15.6 kg; *P*=.61), body fat mass (mean –0.075, SD 1.1 kg; mean –1.0, SD 0.46 kg; *P*=.06), time of MVPA (mean 1.4, SD 7.5 min; mean –1.1, SD 6.9 min; *P*=.32) between the group and control group (Table 2).

**Table 2.** Amount of change before and after intervention between groups[a].

| | Systolic blood pressure | | | Diastolic blood pressure | | |
|---|---|---|---|---|---|---|
| | Estimate | SD | *P* value | Estimate | SD | *P* value |
| Dance | –12.8 | 6.1 | .047 | –9.7 | 3.3 | .006 |
| Sex | –2.8 | 5.9 | .60 | –1.1 | 3.1 | .70 |
| Age | –0.5 | 0.3 | .10 | –0.6 | 0.2 | .001 |
| Pre-MVPA[b] (minutes) | –0.2 | 0.2 | .30 | –0.006 | 0.1 | .09 |

[a]Missing values were excluded from the analysis.

[b]SBP: systolic blood pressure.

[c]DBP: diastolic blood pressure.

[d]MVPA: moderate to vigorous physical activity.

In the multivariate analysis, SBP and DBP improved significantly in the intervention group compared with the control group (mean SBP –12.8, SD 6.1 mm Hg; *P*=.05; mean DBP 9.7, SD 3.3 mm Hg; *P*=.006). For the other covariates, only age showed a significant difference in DBP (*P*=.001; Table 3). No significant harm or unexpected effects were reported during this study.

**Table 3.** Multivariable analysis of systolic/diastolic blood pressure and each response variable[a].

| | Systolic blood pressure | | | Diastolic blood pressure | | |
|---|---|---|---|---|---|---|
| | Estimate | SD | *P* value | Estimate | SD | *P* value |
| Dance | –12.8 | 6.1 | .047 | –9.7 | 3.3 | .006 |
| Sex | –2.8 | 5.9 | .60 | –1.1 | 3.1 | .70 |
| Age | –0.5 | 0.3 | .10 | –0.6 | 0.2 | .001 |
| Pre-MVPA[b] (minutes) | –0.2 | 0.2 | .30 | –0.006 | 0.1 | .09 |

[a]Missing values were excluded from the analysis.

[b]MVPA: moderate to vigorous physical activity.

## Discussion

### Principal Findings

Our results confirmed that regular exercise therapy using dance videos can lower the BP of patients with hypertension, even without monitoring. To the best of our knowledge, this is the first report of this finding.

BP control is crucial to maintaining health. However, various barriers, such as environmental and time constraints, prevent patients from engaging in exercise, which is a useful nonpharmacological therapy for BP control.

### The Relationship Between Exercise and BP

Regarding the relationship between exercise and BP, the antihypertensive effects of aerobic exercise have been well documented in numerous meta-analyses [8,39,40]. Aerobic exercise can significantly decrease SBP and DBP, with specific reductions observed in postmenopausal women and those who participate in combined aerobic and resistance exercises [41]. The American College of Cardiology/American Heart Association guidelines report that exercise therapy can reduce SBP by 2-5 mm Hg and DBP by 1-4 mm Hg [42]. An 8-week stepping exercise program lowered SBP/DBP by 13.1/14.8 mm Hg in older women with stage 1 hypertension [43]. In another study, swimming reduced SBP and DBP by 9 mm Hg over 20 weeks [44]. A meta-analysis of 22 trials (736 participants) examining the effects of regular running on resting BP showed a significant reduction in hypertensive patients' resting BP, with a weighted MD of SBP –5.6 mm Hg (95% CI –9.1 to –2.1; *P*=.01) and DBP –5.2 mm Hg (95% CI –9.0 to –1.4; *P*<.01) [11]. A meta-analysis of 32 studies examining the effects of walking interventions on cardiovascular disease risk factors found a significant improvement in BP among patients with hypertension, with SBP –3.58 mm Hg (95% CI –5.19 to –1.97) and DBP –1.54 mm Hg (95% CI –2.83 to –0.26) [45]. Although the mechanisms underlying these effects are not fully understood, several other factors have been considered. Exercise likely reduces arterial pressure by decreasing cardiac output and total peripheral resistance [46]. Exercise reduces vascular responsiveness to norepinephrine, which increases vascular resistance, and reduces plasma endothelin-1 concentration. Furthermore, endothelium-dependent vasodilation is critically dependent on the production of nitric oxide. Exercise training has been shown to increase nitric oxide production and improve vasodilatory function in healthy participants [47-58]. Vertical head movements during moderate exercise may reduce angiotensin II type 1 receptor expression and BP [59]. Other mechanisms include structural changes in the blood vessels and genetic factors; however, more data are needed [60-62]. In this study, the dance group showed significant improvement in SBP and DBP compared to the control group (mean SBP –12.8, SD 6.1 mm Hg and mean DBP –9.7, SD 3.3 mm Hg). This

improvement is comparable to that observed with other aerobic exercises.

### The Relationship Between Dance and BP and Monitoring Methods in Previous Studies

Dance is a dynamic aerobic endurance exercise that is broadly defined as moving one's body rhythmically to music, usually as a form of artistic or emotional expression. Many health benefits of dance have been realized in recent years. In a previous meta-analysis, the effects of dancing on a large variety of physical health measures were assessed in healthy adults. Studies on healthy adults have found that dance is equal to or greater than exercise in terms of its effectiveness in improving physical health [63-68]. Additionally, a meta-analysis comparing dance with other exercises showed that attrition rates from dance interventions were reported to be lower or equal to exercise, and adherence rates from dance interventions were higher or similar to exercise [21]. In a meta-analysis, dance therapy significantly reduced BP in patients with hypertension, with reductions of approximately 12 mm Hg in SBP and 3.4 mm Hg in DBP [69]. Patients with hypertension undergoing dance movement therapy experience reductions in SBP by 19.2 mm Hg and DBP by 9.5 mm Hg after 4 weeks of twice-weekly sessions [25]. Dances performed in dance movement therapy are often rooted in modern dance [26], but other dance genres also have a positive impact on BP control in patients with hypertension. In aerobic dance, participants saw a decrease in SBP by 18.8 mm Hg and DBP by 8.9 mm Hg over 12 weeks of 45-minute sessions three times a week [27]. Hula dance participants experienced a reduction in SBP by 18.3 mm Hg compared to 7.6 mm Hg in the control group after 12 weeks of 60-minute sessions twice a week [28]. In a study of older adults performing folk dance, SBP decreased from 146.8 mm Hg to 133.8 mm Hg and DBP from 78 mm Hg to 72 mm Hg over 12 weeks of 50-minute sessions three times a week [29]. Additionally, chain dance led to a decrease in SBP by 9 mm Hg and DBP by 6 mm Hg after 6 weeks of 30 to 45-minute sessions twice a week [30]. Overall, dance has been suggested to be highly effective in improving BP, and the results of this study support this.

### Differences Between Previous Dance Studies and Ours

Naturally, exercise prescriptions are meaningless unless implemented by patients. The method of monitoring exercise implementation is likely an important factor in evaluating the effectiveness of exercise therapy in patients with hypertension. In previous studies investigating the relationship between dance exercise prescriptions and BP control, improvements in BP control were observed in all cases. However, as mentioned, in all these studies, the execution of dance exercises was monitored face-to-face or through other means. The most significant difference between this study and the previous research is that we tested the effectiveness of dance-based exercise prescriptions on BP without monitoring. To our knowledge, no previous study has examined the antihypertensive effects of dancing without monitoring. This study is the first to entrust everything to the patients themselves, without monitoring whether the exercise prescriptions were carried out or how accurately the participants performed the dance. In this study, we did not conduct monitoring during the dance sessions; the SBP and DBP in the dance group showed a significant improvement compared with those in the control group. General outpatient care must be carried out in a very short time, lasting only 5-10 minutes, and the existence of a fixed tool that can be used without supervision is thought to be highly effective in the management of lifestyle-related diseases.

Therefore, dance exercises using dance videos may be superior to other forms of exercise in terms of sustainability. Previous noninterventional studies have found that the primary intrinsic motivator for participation in dance was having fun [70] or improving mood [71], whereas participants also experienced significant physical benefits. This was a secondary motivator for initial and maintained participation, thereby likely demonstrating the enjoyment and adherence link that exists in dance. It is presumed that the pleasure and enjoyment experienced by many through dance offers the additional advantage of an increased likelihood of regular participation and adherence, which are essential features for achieving long-term health benefits and could explain the results seen in the included studies. This result is consistent with previous findings. Additionally, in this study, a dance exercise video posted on YouTube was provided as reference material for physical activity. This approach may have facilitated patients' access to an exercise "model," potentially leading to improved adherence to the prescribed physical activity.

### The Significance of Applying This Study to Clinical Medicine

Incorporating exercise prescriptions using YouTube dance exercise videos into outpatient treatment may improve BP control in patients with hypertension, similar to other exercise prescriptions, even in busy and understaffed outpatient settings without monitoring. If video-based dance prescriptions, such as those used in this study, were put into practice, doctors would only need to provide patients with dance prescription videos. This could eliminate the need to spend valuable time during outpatient visits explaining exercises or monitoring exercise routines.

### Limitations

This study had a few limitations.

First, because the patients were recruited from a single university hospital, there may be a risk of selection bias. In the future, this can be improved by recruiting more participants from additional outpatient clinics.

Second, the frequency of dance sessions and the accuracy of movements in the intervention group were unknown. Exercise therapy, intensity, and duration in the control group were also unknown because they were not measured.

Third, the timing of the outpatient visit was generally set at 8 weeks after registration for both BP and body composition measurements; however, there was some variation due to the timing of the outpatient visit.

Fourth, factors such as exercise, diet, and sleep immediately before BP measurement were not standardized because the schedule was adjusted to suit the participants' convenience.

Fifth, since three participants from each group dropped out during the observation period, BP changes in these individuals may have occurred due to antihypertensive medications, making evaluation difficult due to the absence of BP data. Including these participants could reduce the validity of the study; therefore, their exclusion is appropriate.

Despite these limitations, this study remains useful, though it faces constraints due to its focus on verifying the effectiveness of exercise prescriptions through dance videos in outpatient settings.

## Conclusions

This study examined the effects of videos of unsupervised dance exercises on patients with hypertension. The results showed that dance videos were more effective in lowering BP than conventional exercise prescriptions. These results will contribute to exercise therapy for patients with lifestyle-related diseases.

## Acknowledgments

We would like to thank the outpatient nurses and the doctors in charge of the outpatient department of the Department of General Medicine who cooperated with our research. No funding was provided to participants. The equipment used in this research was purchased with research funds from the Department of General Medicine at Juntendo University Hospital. Additionally, some equipment was loaned by the Sportsology Center at the Graduate School of Medicine, Juntendo University. We did not use generative artificial intelligence in our study.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

## Authors' Contributions

MSakairi collected the data. MSakairi analyzed the data and wrote the manuscript with feedback from TM, HT, NY, MSaita, MSuzuki, KF, HF, and TN. TM supervised the project. All authors contributed substantially to the study design and conceptualization, reviewed the manuscript, and approved the final version.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1123 KB - cardio_v9i1e65981_app1.pdf ]

## References

1. GP patient survey 2020. URL: https://www.england.nhs.uk/statistics/2020/07/09/gp-patient-survey-2020/ [accessed 2024-11-28]
2. Yoon SS, Carroll MD, Fryar CD. Hypertension prevalence and control among adults: United States, 2011-2014. NCHS Data Brief 2015(220):1-8 [FREE Full text] [Medline: 26633197]
3. United Nations. World population prospects. 2019. URL: https://population.un.org/wpp/ [accessed 2024-11-28]
4. Arai H, Ouchi Y, Toba K, Endo T, Shimokado K, Tsubota K, et al. Japan as the front-runner of super-aged societies: perspectives from medicine and medical care in Japan. Geriatr Gerontol Int 2015;15(6):673-687 [FREE Full text] [doi: 10.1111/ggi.12450] [Medline: 25656311]
5. Kjeldsen S, Feldman RD, Lisheng L, Mourad J, Chiang C, Zhang W, et al. Updated national and international hypertension guidelines: a review of current recommendations. Drugs 2014;74(17):2033-2051 [FREE Full text] [doi: 10.1007/s40265-014-0306-5] [Medline: 25315030]
6. Pescatello LS, MacDonald HV, Lamberti L, Johnson BT. Exercise for hypertension: a prescription update integrating existing recommendations with emerging research. Curr Hypertens Rep 2015;17(11):87 [FREE Full text] [doi: 10.1007/s11906-015-0600-y] [Medline: 26423529]
7. Umemura S, Arima H, Arima S, Asayama K, Dohi Y, Hirooka Y, et al. The Japanese society of hypertension guidelines for the management of hypertension (JSH 2019). Hypertens Res 2019;42(9):1235-1481. [doi: 10.1038/s41440-019-0284-9] [Medline: 31375757]
8. Börjesson M, Onerup A, Lundqvist S, Dahlöf B. Physical activity and exercise lower blood pressure in individuals with hypertension: narrative review of 27 RCTs. Br J Sports Med 2016;50(6):356-361 [FREE Full text] [doi: 10.1136/bjsports-2015-095786] [Medline: 26787705]
9. Reichert T, Costa RR, Barroso BM, da Rocha VDMB, Delevatti RS, Kruel LFM. Aquatic training in upright position as an alternative to improve blood pressure in adults and elderly: a systematic review and meta-analysis. Sports Med 2018;48(7):1727-1737 [FREE Full text] [doi: 10.1007/s40279-018-0918-0] [Medline: 29651756]

10. Malem R, Ristiani R, Puteh MA. Brisk walking exercise has benefits of lowering blood pressure in hypertension sufferers: a systematic review and meta-analysis. Iran J Public Health 2024;53(4):774-784 [FREE Full text] [doi: 10.18502/ijph.v53i4.15554] [Medline: 39444461]
11. Igarashi Y, Nogami Y. Running to lower resting blood pressure: a systematic review and meta-analysis. Sports Med 2020;50(3):531-541 [FREE Full text] [doi: 10.1007/s40279-019-01209-3] [Medline: 31677122]
12. Clemente FM, Ramirez-Campillo R, Sarmento H. Effects of the small-sided soccer games on blood pressure in untrained hypertensive adults: a systematic review and meta-analysis of randomized controlled trials. Healthcare (Basel) 2021;9(3):345 [FREE Full text] [doi: 10.3390/healthcare9030345] [Medline: 33803787]
13. Igarashi Y, Nogami Y. The effect of regular aquatic exercise on blood pressure: a meta-analysis of randomized controlled trials. Eur J Prev Cardiol 2018;25(2):190-199 [FREE Full text] [doi: 10.1177/2047487317731164] [Medline: 28914562]
14. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
15. The Sasakawa Sports Foundation survey on sports life. 2020. URL: https://www.nippon.com/en/japan-data/h00989/ [accessed 2021-04-22]
16. Vergano LB, Monesi M, Vicenti G, Bizzoca D, Solarino G, Moretti B. Posterior approaches in malleolar fracture: when, why and how. J Biol Regul Homeost Agents 2020;34(3 Suppl. 2):89-95. [Medline: 32856446]
17. Kilgour L, Parker A. Gender, physical activity and fear: women, exercise and the great outdoors. Qual Res Sport Exercise Health 2013;5(1):43-57 [FREE Full text] [doi: 10.1080/2159676x.2012.718619]
18. Lobelo F, de Quevedo IG. The evidence in support of physicians and health care providers as physical activity role models. Am J Lifestyle Med 2016;10(1):36-52 [FREE Full text] [doi: 10.1177/1559827613520120] [Medline: 26213523]
19. The Ministry of Education, Culture, Sports, Science, and Technology, Japan. Course of study guidelines 2012. URL: https://unevoc.unesco.org/home/Explore+the+UNEVOC+Network/centre=288 [accessed 2024-11-28]
20. Amazon. The 2023 SSF National sports-life survey of children and young people. 2020. URL: https://www.amazon.in/National-Sports-Life-Survey-Children-People-ebook/dp/B0D4XRFFK7 [accessed 2024-05-22]
21. Yan AF, Cobley S, Chan C, Pappas E, Nicholson LL, Ward RE, et al. The effectiveness of dance interventions on physical health outcomes compared to other forms of physical activity: a systematic review and meta-analysis. Sports Med 2018;48(4):933-951 [FREE Full text] [doi: 10.1007/s40279-017-0853-5] [Medline: 29270864]
22. Xiong X, Wang P, Li S, Zhang Y, Li X. Effect of baduanjin exercise for hypertension: a systematic review and meta-analysis of randomized controlled trials. Maturitas 2015;80(4):370-378 [FREE Full text] [doi: 10.1016/j.maturitas.2015.01.002] [Medline: 25636242]
23. Peng Y, Su Y, Wang Y, Yuan L, Wang R, Dai J. Effects of regular dance therapy intervention on blood pressure in hypertension individuals: a systematic review and meta-analysis. J Sports Med Phys Fitness 2021;61(2):301-309 [FREE Full text] [doi: 10.23736/s0022-4707.20.11088-0]
24. Yan ZW, Yang Z, Yang JH, Song CL, Zhao Z, Gao Y. Comparison between tai chi and square dance on the antihypertensive effect and cardiovascular disease risk factors in patients with essential hypertension: a 12-week randomized controlled trial. J Sports Med Phys Fitness 2022;62(11):1568-1575 [FREE Full text] [doi: 10.23736/s0022-4707.22.13424-9]
25. Aweto HA, Owoeye OBA, Akinbo SRA, Onabajo AA. Effects of dance movement therapy on selected cardiovascular parameters and estimated maximum oxygen consumption in hypertensive patients. Nig Q J Hosp Med 2012;22(2):125-129. [Medline: 23175912]
26. Japan Dance Therapy Association. What is dance movement therapy?. URL: https://tinyurl.com/2hdw65rh [accessed 2024-11-28]
27. Maruf FA, Akinpelu AO, Salako BL. A randomized controlled trial of the effects of aerobic dance training on blood lipids among individuals with hypertension on a thiazide. High Blood Press Cardiovasc Prev 2014;21(4):275-283 [FREE Full text] [doi: 10.1007/s40292-014-0063-2] [Medline: 24956970]
28. Kaholokula JK, Look M, Mabellos T, Zhang G, de Silva M, Yoshimura S, et al. Cultural dance program improves hypertension management for native Hawaiians and pacific islanders: a pilot randomized trial. J Racial Ethn Health Disparities 2017;4(1):35-46 [FREE Full text] [doi: 10.1007/s40615-015-0198-4] [Medline: 27294768]
29. Rodziewicz-Flis EA, Kawa M, Kaczor JJ, Szaro-Truchan M, Flis DJ, Lombardi G, et al. Changes in selected exerkines concentration post folk-dance training are accompanied by glucose homeostasis and physical performance improvement in older adults. Sci Rep 2023;13(1):8596 [FREE Full text] [doi: 10.1038/s41598-023-35583-w] [Medline: 37237034]
30. Hofgaard J, Ermidis G, Mohr M. Effects of a 6-week Faroese Chain dance programme on postural balance, physical function, and health profile in elderly subjects: a pilot study. Biomed Res Int 2019;2019:5392970 [FREE Full text] [doi: 10.1155/2019/5392970] [Medline: 31392213]
31. Punia S, Kulandaivelan S. Home-based isometric handgrip training on RBP in hypertensive adults—partial preliminary findings from RCT. Physiother Res Int 2020;25(1):e1806. [doi: 10.1002/pri.1806] [Medline: 31418966]
32. The Japan Street Dance Studio Association, a public interest corporation. URL: https://nssa.or.jp/movie/age/ [accessed 2024-11-28]
33. Free BGM Dova-Syndrome. URL: https://dova-s.jp/EN/ [accessed 2024-11-28]

34. Usagawa T, Look M, de Silva M, Stickley C, Kaholokula J, Seto T, et al. Metabolic equivalent determination in the cultural dance of hula. Int J Sports Med 2014;35(5):399-402 [FREE Full text] [doi: 10.1055/s-0033-1353213] [Medline: 24203801]
35. Wu WJ, Yu HB, Tai WH, Zhang R, Hao WY. Validity of actigraph for measuring energy expenditure in healthy adults: a systematic review and meta-analysis. Sensors (Basel) 2023;23(20):8545 [FREE Full text] [doi: 10.3390/s23208545] [Medline: 37896640]
36. Hiroyuki S, Hikihara Y, Okazaki K. Elucidation of activity, sedentary behavior, and sleep patterns that contribute to maintaining weight loss. Descente Sports Sci 2018 May:173-181.
37. Freedson P, Bowles HR, Troiano R, Haskell W. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. Med Sci Sports Exerc 2012;44(1):S1-S4 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399b7e] [Medline: 22157769]
38. Hiroyuki S, Yuki H, Kanzo O. Journal of Epidemiology research on physical activity. Jpn Assoc Exercise Epidemiol 2015;17(1):6-18. [doi: 10.24804/ree.17.6]
39. Dickinson HO, Mason JM, Nicolson DJ, Campbell F, Beyer FR, Cook JV, et al. Lifestyle interventions to reduce raised blood pressure: a systematic review of randomized controlled trials. J Hypertens 2006;24(2):215-233 [FREE Full text] [doi: 10.1097/01.hjh.0000199800.72563.26] [Medline: 16508562]
40. Pescatello LS, MacDonald HV, Ash GI, Lamberti LM, Farquhar WB, Arena R, et al. Assessing the existing professional exercise recommendations for hypertension: a review and recommendations for future research priorities. Mayo Clin Proc 2015;90(6):801-812 [FREE Full text] [doi: 10.1016/j.mayocp.2015.04.008] [Medline: 26046413]
41. Zhou WS, Zheng T, Mao S, Xu H, Wang X, Zhang S. Comparing the effects of different exercises on blood pressure and arterial stiffness in postmenopausal women: a systematic review and meta-analysis. Exp Gerontol 2023;171:111990 [FREE Full text] [doi: 10.1016/j.exger.2022.111990] [Medline: 36397637]
42. Eckel RH, Jakicic JM, Ard JD, de Jesus JM, Miller NH, Hubbard VS, et al. 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. Circulation 2014;129(25):S76-S99 [FREE Full text] [doi: 10.1161/01.cir.0000437740.48606.d1] [Medline: 24222015]
43. Sarinukul C, Janyacharoen T, Donpunha W, Nakmareong S, Ruksapukdee W, Sawanyawisuth K. The effects of stepping exercise on blood pressure, physical performance, and quality of life in female older adults with stage 1 hypertension: a randomized controlled trial. Can Geriatr J 2023;26(1):144-149 [FREE Full text] [doi: 10.5770/cgj.26.632] [Medline: 36865403]
44. Wong A, Kwak YS, Scott SD, Pekas EJ, Son WM, Kim JS, et al. The effects of swimming training on arterial function, muscular strength, and cardiorespiratory capacity in postmenopausal women with stage 2 hypertension. Menopause 2018;26(6):653-658. [doi: 10.1097/GME.0000000000001288] [Medline: 30562322]
45. Murtagh EM, Nichols L, Mohammed MA, Holder R, Nevill AM, Murphy MH. The effect of walking on risk factors for cardiovascular disease: an updated systematic review and meta-analysis of randomised control trials. Prev Med 2015;72:34-43 [FREE Full text] [doi: 10.1016/j.ypmed.2014.12.041] [Medline: 25579505]
46. Abboud FM. The sympathetic system in hypertension. State-of-the-art review. Hypertension 1982;4(3):208-225 [FREE Full text] [doi: 10.1161/01.HYP.4.3_Pt_2.208]
47. Cardillo C, Kilcoyne CM, Waclawiw M, Cannon RO, Panza JA. Role of endothelin in the increased vascular tone of patients with essential hypertension. Hypertension 1999;33(2):753-758 [FREE Full text] [doi: 10.1161/01.hyp.33.2.753] [Medline: 10024340]
48. Schiffrin EL. Endothelin and endothelin antagonists in hypertension. J Hypertens 1998;16(12 Pt 2):1891-1895 [FREE Full text] [doi: 10.1097/00004872-199816121-00007] [Medline: 9886874]
49. Taddei S, Virdis A, Ghiadoni L, Sudano I, Notari M, Salvetti A. Vasoconstriction to endogenous endothelin-1 is increased in the peripheral circulation of patients with essential hypertension. Circulation 1999;100(16):1680-1683 [FREE Full text] [doi: 10.1161/01.cir.100.16.1680] [Medline: 10525485]
50. Maeda S, Tanabe T, Miyauchi T, Otsuki T, Sugawara J, Iemitsu M, et al. Aerobic exercise training reduces plasma endothelin-1 concentration in older women. J Appl Physiol (1985) 2003;95(1):336-341 [FREE Full text] [doi: 10.1152/japplphysiol.01016.2002] [Medline: 12611765]
51. Spier SA, Laughlin MH, Delp MD. Effects of acute and chronic exercise on vasoconstrictor responsiveness of rat abdominal aorta. J Appl Physiol (1985) 1999;87(5):1752-1757 [FREE Full text] [doi: 10.1152/jappl.1999.87.5.1752] [Medline: 10562619]
52. Wiegman DL, Harris PD, Joshua IG, Miller FN. Decreased vascular sensitivity to norepinephrine following exercise training. J Appl Physiol Respir Environ Exerc Physiol 1981;51(2):282-287 [FREE Full text] [doi: 10.1152/jappl.1981.51.2.282] [Medline: 7263435]
53. Maeda S, Miyauchi T, Kakiyama T, Sugawara J, Iemitsu M, Irukayama-Tomobe Y, et al. Effects of exercise training of 8 weeks and detraining on plasma levels of endothelium-derived factors, endothelin-1 and nitric oxide, in healthy young humans. Life Sci 2001;69(9):1005-1016 [FREE Full text] [doi: 10.1016/s0024-3205(01)01192-4] [Medline: 11508642]
54. Kingwell BA. Nitric oxide-mediated metabolic regulation during exercise: effects of training in health and cardiovascular disease. FASEB J 2000;14(12):1685-1696 [FREE Full text] [doi: 10.1096/fj.99-0896rev] [Medline: 10973917]

55. Kingwell BA, Tran B, Cameron JD, Jennings GL, Dart AM. Enhanced vasodilation to acetylcholine in athletes is associated with lower plasma cholesterol. Am J Physiol Heart Circ Physiol 1996;270(6):H2008-H2013 [FREE Full text] [doi: 10.1152/ajpheart.1996.270.6.h2008]
56. Higashi Y, Sasaki S, Kurisu S, Yoshimizu A, Sasaki N, Matsuura H, et al. Regular aerobic exercise augments endothelium-dependent vascular relaxation in normotensive as well as hypertensive subjects: role of endothelium-derived nitric oxide. Circulation 1999;100(11):1194-1202 [FREE Full text] [doi: 10.1161/01.cir.100.11.1194] [Medline: 10484540]
57. Higashi Y, Sasaki S, Sasaki N, Nakagawa K, Ueda T, Yoshimizu A, et al. Daily aerobic exercise improves reactive hyperemia in patients with essential hypertension. Hypertension 1999;33(1 Pt 2):591-597 [FREE Full text] [doi: 10.1161/01.hyp.33.1.591] [Medline: 9931171]
58. Green DJ, Maiorana A, O'Driscoll G, Taylor R. Effect of exercise training on endothelium-derived nitric oxide function in humans. J Physiol 2004;561(Pt 1):1-25 [FREE Full text] [doi: 10.1113/jphysiol.2004.068197] [Medline: 15375191]
59. Murase S, Sakitani N, Maekawa T, Yoshino D, Takano K, Konno A, et al. Interstitial-fluid shear stresses induced by vertically oscillating head motion lower blood pressure in hypertensive rats and humans. Nat Biomed Eng 2023;7(11):1350-1373 [FREE Full text] [doi: 10.1038/s41551-023-01061-x] [Medline: 37414976]
60. Pescatello LS, Franklin BA, Fagard R, Farquhar WB, Kelley GA, Ray CA. Exercise and hypertension. Med Sci Sports Exerc 2004;36(3):533-553 [FREE Full text] [doi: 10.1249/01.mss.0000115224.88514.3a] [Medline: 15076798]
61. Blanchard BE, Tsongalis GJ, Guidry MA, LaBelle LA, Poulin M, Taylor AL, et al. RAAS polymorphisms alter the acute blood pressure response to aerobic exercise among men with hypertension. Eur J Appl Physiol 2006;97(1):26-33 [FREE Full text] [doi: 10.1007/s00421-006-0142-8] [Medline: 16468060]
62. Santana HA, Moreira SR, Neto WB, Silva CB, Sales MM, Oliveira VN, et al. The higher exercise intensity and the presence of allele I of ACE gene elicit a higher post-exercise blood pressure reduction and nitric oxide release in elderly women: an experimental study. BMC Cardiovasc Disord 2011;11(1):71 [FREE Full text] [doi: 10.1186/1471-2261-11-71] [Medline: 22136292]
63. Barene S, Holtermann A, Oseland H, Brekke O, Krustrup P. Effects on muscle strength, maximal jump height, flexibility and postural sway after soccer and Zumba exercise among female hospital employees: a 9-month randomised controlled trial. J Sports Sci 2016;34(19):1849-1858 [FREE Full text] [doi: 10.1080/02640414.2016.1140906] [Medline: 26849477]
64. Milburn S, Butts NK. A comparison of the training responses to aerobic dance and jogging in college females. Med Sci Sports Exerc 1983;15(6):510-513 [FREE Full text] [doi: 10.1249/00005768-198315060-00012]
65. Barene S, Krustrup P, Brekke OL, Holtermann A. Soccer and Zumba as health-promoting activities among female hospital employees: a 40-weeks cluster randomised intervention study. J Sports Sci 2014;32(16):1539-1549 [FREE Full text] [doi: 10.1080/02640414.2014.906043]
66. Rani SU, Singh SP. Effect of selected yogic practices and aerobic dance on health related physical fitness variables among nicobari women students. IJPEFS 2013;2(4):76-79 [FREE Full text] [doi: 10.26524/13420]
67. Garber CE, McKinney JS, Carleton RA. Is aerobic dance an effective alternative to walk-jog exercise training? J Sports Med Phys Fitness 1992;32(2):136-141. [Medline: 1434581]
68. Isler AK, Koşar SN, Korkusuz F. Effects of step aerobics and aerobic dancing on serum lipids and lipoproteins. J Sports Med Phys Fitness 2001;41(3):380-385. [Medline: 11533570]
69. Conceição LSR, Neto MG, do Amaral MAS, Martins-Filho PRS, Carvalho VO. Effect of dance therapy on blood pressure and exercise capacity of individuals with hypertension: a systematic review and meta-analysis. Int J Cardiol 2016;220:553-557 [FREE Full text] [doi: 10.1016/j.ijcard.2016.06.182] [Medline: 27390986]
70. Maraz A, Király O, Urbán R, Griffiths MD, Demetrovics Z. Why do you dance? Development of the dance motivation inventory (DMI). PLoS One 2015;10(3):e0122866 [FREE Full text] [doi: 10.1371/journal.pone.0122866] [Medline: 25803301]
71. Lakes KD, Marvin S, Rowley J, Nicolas MS, Arastoo S, Viray L, et al. Dancer perceptions of the cognitive, social, emotional, and physical benefits of modern styles of partnered dancing. Complement Ther Med 2016;26:117-122 [FREE Full text] [doi: 10.1016/j.ctim.2016.03.007] [Medline: 27261991]


## Abbreviations

**BP:** blood pressure
**CONSORT:** Consolidated Standards of Reporting Trials
**DBP:** diastolic blood pressure
**MD:** mean difference
**MET:** metabolic equivalent of task
**MVPA:** moderate to vigorous physical activity
**SBP:** systolic blood pressure
**UMIN:** University Hospital Medical Information Network

*Edited by A Coristine; submitted 31.08.24; peer-reviewed by H Shah, T Akindahunsi; comments to author 30.09.24; revised version received 25.11.24; accepted 25.11.24; published 09.01.25.*

*Please cite as:*
*Sakairi M, Miyagami T, Tabata H, Yanagisawa N, Saita M, Suzuki M, Fujibayashi K, Fukuda H, Naito T*
*Efficacy of Unsupervised YouTube Dance Exercise for Patients With Hypertension: Randomized Controlled Trial*
*JMIR Cardio 2025;9:e65981*
*URL: https://cardio.jmir.org/2025/1/e65981*
*doi:10.2196/65981*
*PMID:*

Original Paper

# Exploring Stakeholder Perspectives on the Barriers and Facilitators of Implementing Digital Technologies for Heart Disease Diagnosis: Qualitative Study

Kamilla Abdullayev[1], MSc; Tim J A Chico[2], PhD; Jiana Canson[1]; Matthew Mantelow[3], PhD; Oli Buckley[4], PhD; Joan Condell[3], PhD; Richard J Van Arkel[5], PhD; Vanessa Diaz-Zuccarini[6], PhD; Faith Matcham[1], PhD

[1]School of Psychology, University of Sussex, Falmer, United Kingdom

[2]Clinical Medicine, School of Medicine and Population Health, The Medical School, University of Sheffield, Sheffield, United Kingdom

[3]School of Computing, Engineering and Intelligent Systems, Ulster University at Magee, Londonderry, United Kingdom

[4]Department of Computer Science, Loughborough University, Loughborough, United Kingdom

[5]Department of Mechanical Engineering, Imperial College London, London, United Kingdom

[6]Department of Mechanical Engineering, University College London, London, United Kingdom

**Corresponding Author:**
Faith Matcham, PhD
School of Psychology
University of Sussex
Pevensey 1
Falmer, BN1 9QH
United Kingdom
Phone: 44 01273 606755
Email: f.matcham@sussex.ac.uk

## Abstract

**Background:** Digital technologies are increasingly being implemented in health care to improve the quality and efficiency of care for patients. However, the rapid adoption of health technologies over the last 5 years has failed to adequately consider patient and clinician needs, which results in ineffective implementation. There is also a lack of consideration for the differences between patient and clinician needs, resulting in overgeneralized approaches to the implementation and use of digital health technologies.

**Objective:** This study aimed to explore barriers and facilitators of the implementation of digital technologies in the diagnosis of heart disease for both patients and clinicians, and to provide recommendations to increase the acceptability of novel health technologies.

**Methods:** We recruited 32 participants from across the United Kingdom, including 23 (72%) individuals with lived experience of heart disease and 9 (28%) clinicians involved in diagnosing heart disease. Participants with experience of living with heart disease took part in semistructured focused groups, while clinicians contributed to one-to-one semistructured interviews. Inductive thematic analysis using a phenomenological approach was conducted to analyze the resulting qualitative data and to identify themes. Results were discussed with a cardiovascular patient advisory group to enhance the rigor of our interpretation of the data.

**Results:** Emerging themes were separated into facilitators and barriers and categorized into resource-, technology-, and user-related themes. Resource-related barriers and facilitators related to concerns around increased clinician workload, the high cost of digital technologies, and systemic limitations within health care systems such as outdated equipment and limited support. Technology-related barriers and facilitators included themes related to reliability, accuracy, safety parameters, data security, ease of use, and personalization, all of which can impact engagement and trust with digital technologies. Finally, the most prominent themes were the user-related barriers and facilitators, which encompassed user attitudes, individual-level variation in preferences and capabilities, and impact on quality of health care experiences. This theme captured a wide variety of perspectives among the sample and revealed how patient and clinician attitudes and personal experiences substantially impact engagement with digital health technologies across the cardiovascular care pathway.

**Conclusions:** Our findings highlight the importance of considering both patient and clinician needs and preferences when investigating the barriers and facilitators to effective implementation of digital health technologies. Facilitators to technology

adoption include the need for cost-effective, accurate, reliable, and easy-to-use systems as well as adequate setup support and personalization to meet individual needs. Positive user attitudes, perceived improvement in care quality, and increased involvement in the care process also enhance engagement. While both clinicians and patients acknowledge the potential benefits of digital technologies, effective implementation hinges on addressing these barriers and leveraging facilitators to ensure that the technologies are perceived as useful, safe, and supportive of health care outcomes.

**International Registered Report Identifier (IRRID):** RR2-10.1136/bmjopen-2023-072952

***(JMIR Cardio 2025;9:e66464)*** doi:10.2196/66464

**KEYWORDS**

heart disease; digital technologies; stakeholder perspectives; qualitative research; digital technology; health technology; heart; cardio; cardiology; cardiovascular; qualitative; focused group; quality of care; efficiency; digital health; mobile phone; artificial intelligence; AI

## Introduction

### Background

There has been a sharp rise in the use of digital health technologies in health care, particularly after the COVID-19 pandemic, which drove rapid adoption of remote measurement and consultation technologies [1-3]. In parallel, there has been a rapid growth in the use of consumer *well-being* devices marketed directly to citizens that monitor a range of health measures, such as sleep and heart rate [1-3]. Cardiovascular medicine has been one of the earliest adopters of digital technology in health care because aspects of cardiovascular health, such as electrocardiograms (ECGs), are already proven to be clinically relevant and are measurable using both medical devices and consumer wearables [4-6].

The potential benefits of using digital health technologies within cardiovascular health care are considerable, including early identification and modification of risk factors such as diabetes or hypertension; earlier, faster, or more accurate diagnosis; personalized treatment and management plans; improved ability to monitor disease and detect deterioration; and improved symptom assessment [7]. Meanwhile, health care systems are facing increasing challenges in delivering services designed in a predigital era. Existing care pathways remain rooted in face-to-face clinical assessments and siloed data about the patient across different analog and digital systems that are inaccessible to both the patient and their different care teams.

Digital health technologies could help address factors that contribute to delayed or inaccurate diagnosis of cardiovascular diseases [8]. An example of such an emerging technology is digital twins, which uses mathematical models to process data that are continuously updated to monitor various physiological symptoms over time [9-11]. This allows for the capture of longitudinal symptom data, provides customizable feedback for patients to help them alter behavior and self-manage their condition, and improves patient-clinician communication [12]. This efficient processing of large amounts of cardiovascular data highlights the substantial cost benefits of implementing digital health technologies [13].

The potential of digital technologies to improve health care has often been discussed, particularly by policy makers. However, it is also important to acknowledge that these novel technologies may pose risk, have negative effects on the users and the health care system, or face resistance from patients and clinicians. During the COVID-19 pandemic, patients reported several barriers to engagement with telehealth, including the lack of *human* contact, concerns related to confidentiality and data security, and a requirement for training in the use of new platforms [3]. Several qualitative studies have examined technology engagement among patients with cardiovascular diseases [14,15]. One recent review revealed 4 interrelated themes across 7 qualitative studies, including trust, safety and confidence, functionality and affordability, and risks and assurance, highlighting the complexity of factors contributing to patient engagement [14]. However, the focus of previous investigations has been primarily on technology used in rehabilitation or self-management of the confirmed disease [14,16-19]. However, the most common first stage of medical care is the diagnosis of symptoms that may reflect underlying heart disease, with an estimated 39% of adults experiencing symptoms that can reflect possible underlying heart disease such as chest pain [20]. Therefore, the initial onset of symptoms that may indicate cardiovascular problems affects a far greater number of people than those dealing with recovery from or management of heart disease. Furthermore, the diagnosis stage often comes with increased stress, frustration, and confusion for the patient and their families [21,22]. Thus, specific research is needed to understand the factors that influence the uptake of digital technologies at the stage of diagnosis, as these factors may differ from those that influence the use of technologies in people with proven heart disease.

Moreover, there is rarely a combined focus on both clinician and patient views, which prevents our ability to capture a more holistic perspective on the implementation of health care technology in clinical settings. Patients and clinicians have different needs and expectations of digital technologies, requiring specific exploration of approaches that can address these needs and expectations simultaneously. Al-Naher et al [23] examined factors influencing engagement in remote health care in heart failure and included both patient and clinician perspectives in their review. However, their final conclusions did not differentiate between these different user groups, applying the resulting 5 overarching themes (convenience, ease of use, education, clinical care, and communication) to both groups to provide insight to improve engagement [23], without adjustment based on user-specific needs. Meanwhile, 1 scoping review on the uptake of digital health technology across

cardiovascular care provided separate barriers and facilitators between patient-level and clinician-level perspectives [24]. Their findings suggest that specific considerations should be made regarding user needs when attempting to implement acceptable and useful digital health technologies across different stages of cardiovascular care.

Ultimately, there remains a substantial gap in our understanding of the factors impacting engagement with digital health technologies for heart disease diagnosis across patients and clinicians. Therefore, more work is needed to provide stakeholder-led insights into specific barriers to target and facilitators to consider in the early stages of novel technology development, to improve engagement with, and thus the efficacy of, novel digital health technologies aiming to improve the accuracy and efficiency of heart disease diagnosis.

### Objectives

We used a qualitative approach to address the following objectives:

- Understand patients' and clinicians' views on the barriers and facilitators to the implementation of digital technologies for the diagnosis of heart disease
- Explore whether these perspectives on digital technology differ between patients and clinicians
- Provide evidence-based design considerations for novel digital health technologies to allow for more effective implementation for the diagnosis of heart disease

## Methods

### Overview

Our protocol and methodology have been previously published [25]. This study was conducted as part of a wider project aiming to test technologies available to diagnose a range of heart diseases and establish the most useful ways of communicating data back to clinicians and patients. The findings from this work have contributed to the development of testing priorities and procedures for a larger quantitative trial. The project represents a collaboration between clinical and research institutions across the United Kingdom.

The study was conducted and reported according to COREQ (Consolidated Criteria for Reporting Qualitative Research) [26] guidelines. The question topic guide involved 2 main parts: experiences relating to diagnostic delays and errors, and investigation of barriers and facilitators of engagement with technologies throughout the heart disease diagnosis pathway (Multimedia Appendix 1).

We have previously reported stakeholder experiences of heart disease diagnosis, specifically aiming to identify challenges contributing to delayed and inaccurate diagnosis [12]. This paper presents additional data collected to identify barriers and facilitators to the implementation of digital technologies for heart disease diagnosis, which are critical for uptake into clinical care.

### Study Design

A qualitative approach was taken to capture the depth and complexity of technology-related challenges faced by both patients and clinicians. We conducted semistructured focus groups with people with lived experience (LE) of heart disease to facilitate discussions on shared perspectives regarding the use of digital health technologies and to allow for direct comparisons among a range of diverse experiences with technology, which may have been missed in a one-on-one interview.

We conducted 1:1 interviews with clinicians to allow greater flexibility around their schedules and collect information across a range of clinical specialties.

### Patient and Public Involvement

All participant-facing materials were reviewed by a Sheffield-based cardiovascular patient advisory group. This ensured the information sheet, consent form, and focus group topic guides were accessible and easy to understand, including any technology-related terminology used. This led to the inclusion of a detailed description of the meaning of *digital*, followed by several examples of digital technologies throughout the questions covered.

### Study Population

Inclusion criteria for LE participants were a previous diagnosis of heart disease, aged ≥18 years, able to speak English sufficiently for participation, and able to consent to participate. Exclusion criteria included major cognitive impairment or dementia preventing participation. The inclusion criteria for clinicians were >6 months of experience in the diagnosis of heart disease, aged ≥18 years, able to speak English, and able to consent to participation.

The number of participants recruited for focus groups and interviews was based on pragmatic considerations [27], such as the time available for data collection against the wider project deadlines and the research team's previous experience conducting qualitative research with clinicians [25]. With these practical considerations alongside recent evidence that data saturation can be achieved in as little as 9 interviews and 4 focus groups [28], we aimed to recruit between 4 and 6 LE participants across 4 focus groups to allow adequate time for each participant to share their views and experiences, and to interview 10 clinicians to achieve data saturation.

### Procedure

All participants were recruited in the United Kingdom, and data were collected between November 2022 and April 2023. We implemented a decentralized recruitment strategy, recruiting LE participants via Prolific (a web-based research platform), a panel for patients with cardiovascular diseases at the Sheffield University, and from UK-based participants from the Remote Assessment of Disease and Relapse–Major Depressive Disorder research study who had consented to be contacted for future research purposes [29]. Study information sheets were sent to people identified as meeting the eligibility criteria, with the advice to contact the study team if they were interested in participating. Study details were additionally shared on X,

formerly known as Twitter. Individuals interested in participating were contacted via email to arrange an introductory phone call to confirm interest and eligibility. In this meeting, FM described the research and the procedure of the study. Recruitment materials can be found in Multimedia Appendix 2.

Clinicians were recruited using purposive sampling via personal and professional connections and a registered general physician Facebook (Meta Platforms, Inc) group. The study information sheets were posted on the Facebook group, with interested clinicians advised to contact the study team directly. Among them, clinicians represent a range of clinical roles across the heart disease pathway, from diagnosis through to long-term management. However, for the purposes of this study, we exclusively recruited those who diagnose heart disease on a regular basis. All information was given to clinicians via email before the web-based interview.

Consent and baseline demographic data were collected via web-based Qualtrics (Qualtrics International, Inc) surveys before qualitative data collection (Multimedia Appendix 3). The focus groups and interviews follow a preapproved, semistructured question schedule. Each focus group included either 5 or 6 participants. All focus groups and interviews were conducted on the web using Zoom (Zoom Video Communications), with focus groups lasting about 90 minutes and interviews ranging between 30 and 90 minutes, based on clinician availability. Interviews and focus groups were facilitated by KA, a psychology graduate working full time on the project. KA had no ongoing relationship with the participants and was not involved in their clinical care. She had neither previous experience in cardiology nor assumptions or expectations of the data. To support participants who may have found it challenging to engage with general questions about barriers and facilitators for digital technologies as a broad category, we included follow-up prompts and clarifying examples to help participants contextualize their responses, for instance, the provision of specific scenarios or requests to reflect on their experiences with technologies such as wearables, portable ECG monitors, or smartphones.

## Ethical Considerations

This study was reviewed and approved by the Sciences & Technology Cross-School Research Ethics Council at the University of Sussex (reference ER/FM409/1). It was conducted according to institutional and international guidelines for ethical research practices and complies with the Declaration of Helsinki regulations. Informed consent for each participant was acquired before data collection. Participants were provided with detailed information about the study objectives, procedures, and rights, including the right to withdraw at any time without penalty. The privacy and confidentiality of all participants was safeguarded through strict data protection measures. The focus group and interviews were audio recorded, anonymized, and then transcribed verbatim before analysis, with encryption and secure storage protocols implemented to prevent unauthorized data access. Field notes made during the focus groups were destroyed once transcripts were deidentified and finalized. Participants were compensated for their time with a £25 (US $31) Amazon voucher.

## Data Analysis

Data relating to patient and clinician perspectives on the facilitators and barriers of effective implementation of digital technologies into heart disease diagnosis were included in this analysis. Sample sociodemographic characteristics were also collected.

We conducted an inductive thematic analysis using a phenomenological approach, as this allowed us to be led by the data when exploring emerging themes related to stakeholder experiences. Our method was characteristic of a small q approach, as we followed the postpositivist framework of qualitative analysis to ensure the reliability of the resulting themes related to stakeholder experiences of heart disease diagnosis [30]. KA used NVivo (Lumivero) to conduct the first round of analysis, following the steps recommended by Braun and Clarke [31]. We used the 6-phase approach outlined by Braun and Clarke [31] to identify, analyze, and report patterns (themes) within the data. The six phases included the following: (1) familiarization with the data through reading and rereading, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) writing the report.

## Reflexivity and Positionality

To ensure methodological rigor, we adhered to the best practices outlined by Braun and Clarke [30], particularly focusing on avoiding common problems in thematic analysis, such as insufficient reflexivity or unclear connections between data and themes. In line with this updated guidance, we paid particular attention to how our own assumptions and positionalities might have influenced the analysis process. This reflexive approach was an integral part of our analysis, and we constantly questioned how our perspectives as researchers may have shaped the interpretation of the data.

We remained mindful of power dynamics, particularly during the clinician interviews and patient focus groups. Our familiarity with the clinical context and our personal experiences in conducting qualitative research shaped the way we interacted with participants and interpreted their responses. We also reflected on how the context of data collection (focus group vs individual interview) may influence the themes arising from the data and acknowledged and discussed these throughout the analysis process. This reflexive stance was crucial to ensure that we did not impose our own perspectives on the data, and we actively engaged in discussions with colleagues to challenge potential biases and enhance the trustworthiness of our findings.

## Scientific Rigor

We applied several strategies to ensure the trustworthiness of the study, addressing the dimensions of confirmability, dependability, credibility, and transferability.

To enhance confirmability, we maintained an audit trail throughout the study, documenting each step of the data collection and analysis process. This included detailed notes on our analytical decisions and the rationale for theme development.

We ensured dependability by using a consistent approach to data collection, using semistructured interview guides, and by providing clear descriptions of the process of data analysis. Any deviations from the original plan were noted, and we made sure that the methods were applied systematically across all participants.

Credibility was enhanced through member checking, where we invited participants and other experts by experience to review and comment on the emerging findings. This process allowed us to verify our interpretations and ensure that they accurately represented participants' experiences and perspectives. This was achieved through presenting the results of the first round of thematic analysis, which were presented to clinicians in the form of a research poster at the British Cardiology Society conference to increase the transferability of our results to a wider sample. A QR code was provided next to the poster, allowing clinicians to scan it and provide their reflections on whether we captured their experiences or comment on what was missing. Those unable to scan the code (eg, did not have a mobile available on hand) provided verbal feedback to the research poster presenter (KA). Feedback from 5 clinicians was integrated into the later stages of analysis.

We also consulted with a Sheffield-based cardiovascular patient advisory group again to provide further insight on the results of our analysis. Preliminary results were presented via a series of presentation slides summarizing the key themes that emerged. Verbal discussions were facilitated by the lead researcher (KA), and the meeting minutes were written up by JC.

# Results

## Sample Demographics

In total, 4 patient focus groups (n=23) and 9 individual clinician interviews were performed (n=32), shown in Figure 1. This represents 21.8% (32/147) of individuals initially contacted and 65% (32/49) of individuals who expressed initial interest in taking part. The sample of this study is reported in Table 1. This is the same group of participants that was used in the study by Abdullayev et al [12]; therefore, participants' demographics are the same.

**Figure 1.** A flowchart of recruitment of participants, from initial contact to analysis. RADAR-MDD: Remote Assessment of Disease and Relapse–Major Depressive Disorder.

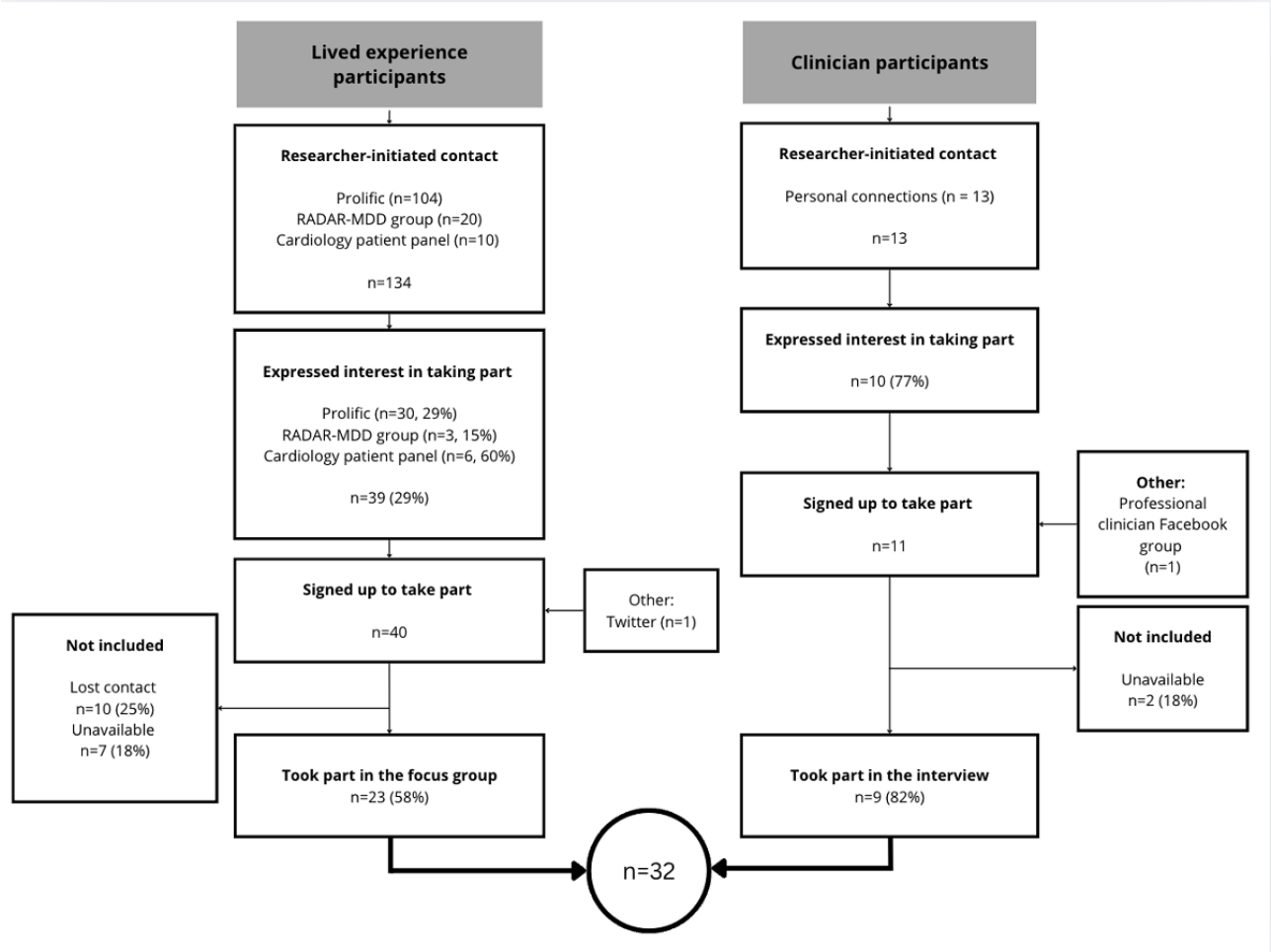


**Table 1.** Demographic characteristics of the sample (n=32).

| Characteristic | Total sample (n=32) | LE[a] participants (n=23) | Clinician (n=9) |
|---|---|---|---|
| Age (y), mean (SD; range) | 58.0 (12.2; 31-76) | 61.3 (11.5; 31-76) | 48.5 (9.1; 35-60) |
| **Sex, n (%)** | | | |
| Male | 22 (69) | 16 (70) | 6 (67) |
| Female | 10 (31) | 7 (30) | 3 (33) |
| **Race and ethnicity, n (%)** | | | |
| Asian | 4 (12) | 2 (9) | 2 (22) |
| Black | 0 (0) | 0 (0) | 0 (0) |
| White | 27 (84) | 21 (91) | 6 (67) |
| Other (Arab) | 1 (3) | 0 (0) | 1 (11) |
| **Income bracket, n (%)** | | | |
| <£15,000 (<US $18,800) | 6 (19) | 6 (26) | 0 (0) |
| £15,000-£24,000 (US $18,800-US $30,200) | 4 (12) | 4 (17) | 0 (0) |
| £24,000-£40,000 (US $30,200-US $50,300) | 8 (25) | 7 (30) | 1 (11) |
| £40,000-£55,000 (US $50,300-US $69,200) | 5 (16) | 5 (22) | 0 (0) |
| >£55,000 (>US $69,200) | 7 (22) | 1 (4) | 6 (67) |
| Not disclosed | 2 (6) | 0 (0) | 2 (22) |

[a]LE: lived experience.

Most clinicians (6/9, 67%) had been in practice for >20 years, representing primary (4/9, 44%), secondary (4/9, 44%), and emergency (1/9, 11%) care services. Most of the clinicians (8/9, 89%) reported feeling fairly to very confident using digital technologies, compared to 70% (16/23) of LE participants. All participants used at least these 3 devices: televisions, mobile phones, and laptops. The majority (27/32, 84%) also reported regularly using tablets or desktop computers. Table 1 summarizes the demographic and clinical characteristics of the sample.

## Analysis Results

Our analysis identified 6 themes arising from the participants' views on digital technologies for the diagnosis of heart disease. A review of our efforts to increase the transferability of our findings via discussions with the Patient Advisory Board and clinicians attending a cardiology conference confirmed the value of considering both clinician and patient perspectives, as they felt this was key to implementing novel technology into health care. Insights provided by the advisory group reinforced confidence that our data fully captured the experience of stakeholders and resonated with their own LE.

Neither form of cross validation resulted in major changes to the analysis; however, it supported the organization and description of the themes and subthemes reported. While it is not possible to remove the subjective bias of the researchers conducting the analysis, this patient and public involvement–led approach to thematic analysis increases the credibility of our findings, which ultimately increases its transferability beyond our sample.

We organized these 6 themes into 2 key categories: barriers (defined as factors that prevent effective implementation) and facilitators (ways to enhance engagement among stakeholders). Textbox 1 summarizes the organization of the 6 themes that emerged from the data.

**Textbox 1.** Summary of the 6 themes emerging from the results of a thematic analysis with a phenomenological approach.

**Themes and subthemes**

- Resource-related barriers: clinician workload, cost implications, and systemic barriers
- Technology-related barriers: complexity of technology, data security and privacy issues, safety concerns, and unreliability
- User-related barriers: negative user attitudes, worsening care experience, and individual-level variation
- Resource-related facilitators: cost-effectiveness, efficiency, and setup support
- Technology-related facilitators: accuracy and reliability, adequate safety considerations, ease of use, patients' right to data, and personalization
- User-related facilitators: adapting to individual characteristics, positive user attitudes, and improving quality of care experience

## Theme 1: Resource-Related Barriers

### *Digital Technologies Can Add to Clinician Workload*

Several clinicians raised considerable concerns regarding additional workload resulting from novel digital technologies being implemented into diagnosis. These participants emphasized that this would be a substantial barrier to the uptake of such health technologies given the current resource restraints within the National Health Service (NHS). Such concerns were not present among patient perspectives:

> *If it was going to make more work for me, if it was...to create any hassle for me I'm not interested.* [Clinician8; male; aged 52 years]

### *Digital Technologies Come With Cost Implications*

Another resource-related barrier was the potential costs of digital technologies, both for the individual and the health care system. Clinicians highlighted current issues related to an imbalance between the cost versus benefits of collecting more patient health data and using it to improve patient health outcomes:

> *At best [they] had only marginal health, marginal impact but the cost of gathering the data and retrieving the important ones proved to be enormous.* [Clinician1; male; aged 60 years]

Patient perspectives also acknowledged how resource limitations within health care systems present challenges with implementing novel technologies in a sustainable way, as there appears to be a lack of connection between the development versus the implementation of digital health solutions:

> *That is what happens in the NHS. They all go off, do something, invent something and never do, they all come together because it costs billions of pounds to do it.* [LE17; male; aged 65 years]

### *Digital Technologies Are Not Immune to Systemic Barriers*

Both clinicians and patients described how existing systemic barriers would prevent effective implementation due to a lack of access to appointments or equipment, a lack of support in initial setup, and difficulties integrating novel technologies into outdated NHS systems. Clinicians expressed doubt in their ability to support patients in setting up a device to aid with diagnosis within the limited appointment time they currently have:

> *GP appointments are 10 to 15 minutes, so how long is it going to take to explain this app, and how it works to them, and expect them to fill it in?* [Clinician2; female; aged 38 years]

Patients also shared frustrations with how outdated technology is within the NHS and how this inevitably acts as a barrier to the implementation of new technologies that could be used to improve heart disease diagnosis:

> *Sadly, the NHS is about 20 years behind with technology for a whole host of reasons.* [LE17; male; aged 65 years]

## Theme 2: Technology-Related Barriers

### *Complexity of Technology*

The complexity of novel technology appears to be an important factor in engagement, as anything with too many steps or too many features to be learned will demotivate an individual's engagement and produce inaccurate or incomplete data, which clinicians will not be able to use. Clinicians described how the complexity of a device will determine their willingness to engage with novel technologies:

> *I think how long or how easy or difficult it is to put or use this device, set it up and have it running and showing a patient what's involved.* [Clinician13; male; aged 49 years]

Patients echoed these concerns, highlighting how increased complexity results in more errors within the data and prevents people from engaging with the device or program:

> *I think that the more complex it is, the more there is room for error, for a start, of actually producing the wrong data. And the second thing is that it may actually discourage people from using it.* [LE29; male; aged 73 years]

### *Issues With Data Security and Privacy*

A key concern related to technology was the way sensitive health data would be protected. Clinicians reflected on potential issues that would arise if patients were not assured that their health data were being handled appropriately:

> *I can see some problems that include confidentiality, you know, these are personal information so you know we just have to make sure it's very secure and you don't know who has got access to this to this information.* [Clinician7; male; aged 44 years]

This concern was also seen among patient perspectives, with fears of large corporations having access to their health data acting as barriers to engaging in health technologies:

> *I'm not too sure whether they should be making money out of people's illnesses or symptoms. I suppose it's the data protection aspect of it.* [LE4; male; aged 76 years]

### *Concerns With Safety*

Given the risks associated with monitoring symptoms before diagnosis, concerns related to the safety of the patient presented as an important barrier for both clinician and patient engagement. Clinicians emphasized the risks associated with collecting health data to monitor symptoms due to difficulties related to establishing safety parameters within the monitoring devices:

> *I think there is a governance issue about asking patients a question and then not processing safely the answer, to safety net them and the challenge there is getting the balance of safety versus being, you know, setting the threshold for seeking extra help to them and that's where I think we've really struggled and*

> *never quite got it right.* [Clinician8; male; aged 52 years]

Moreover, patients expressed feelings of being unsafe in the case of emergency situations when their symptoms are being monitored remotely and doubt that health care staff would respond appropriately if their health was deemed at risk by the technology:

> *My worry about this is quite simple that the system would work but nobody would pick up on it, or actually do something about it if some if there was an emergency.* [LE11; male; aged 70 years]

### *Unreliability of Health Technologies*

In addition to safety concerns, potential unreliability of a technology also emerged as a potential barrier to engagement. Clinicians described situations where they would be reluctant to depend on technology, as they do not feel confident in the reliability of the information it relays to the health care staff:

> *So to say to me, somebody's got a heart attack when they haven't, yeah, it's massive. So I'm not suggesting that AI is doing that all the time, right, left, and centre. It's definitely not doing that but it can do that.* [Clinician1; male; aged 60 years]

Similarly, patients shared doubts regarding how much they would be willing to rely on technological devices due to practical liabilities such as internet connection failure or poor connection in particular regions, as they fear it would pose a greater risk to their health compared to traditional approaches:

> *Another concern that comes to mind is how reliable it is in terms of the you know we're all used to the internet going down like you lost your Internet connection, that could affect the technology used in this area. What happens if it all goes down, because what's the back up? That's a very valid concern.* [LE19; male; aged 64 years]

## Theme 3: User-Related Barriers

### *The Power of Negative User Attitudes*

Negative attitudes toward the use of digital technology within health care were recognized as a potential barrier to engagement in several ways. First, distrust of technology providing reliable and useful information was evident among clinicians, highlighting how user attitudes might be influencing the way novel technologies are being implemented:

> *The blanket belief in AI is rubbish and AI can come up with rubbish if you are not careful.* [Clinician1; male; aged 60 years]

Meanwhile, another clinician felt that patients were more likely to possess this deep-rooted distrust in technology, suggesting there are still fears related to unethical health data collection, storage, and use:

> *Some of these conspiracy type theories where they think that what they're being spied on.* [Clinician12; male; aged 59 years]

Some patients reflected that they would prefer not to have technology involved in the diagnosis pathway. They believed the health care system is implementing these novel systems to save money and do not care about how this impacts patient experiences and quality of care:

> *I just find it, it's an extra barrier we'd rather not have, but because it's cheap, and that doesn't feel great to be treated in a cheap way, but that's what it's come down to, I think, which is very sad.* [LE28; female; aged 50 years]

Finally, a particularly influential user attitude is related to how useful or effective technology solutions were perceived to be. Both patients and clinicians reflected that they would not use a technology if they believed it was not going to benefit them or their patient. This highlights how refusing to engage in technology can be a rational decision made by the user, based on their personal beliefs regarding the potential utility:

> *There's no point...if you get them to record stuff and cardiology don't want it, and don't look at it then actually they're not going to use it.* [Clinician2; female; aged 38 years]

> *Why a chat bot when you can ring 111, and get the same advice from an actual living person?* [LE5; female; aged 61 years]

### *They Worsen Our Care Experience*

Another barrier to engagement was the belief that the use of digital technology would worsen the quality of care. The burden of excessive interaction emerged as a potential barrier to engagement, as patients reflected on how frustration resulted in disengagement when patients are expected to dedicate a lot of their time to input data and track their symptoms:

> *I think the interactions got to be quite, quite minimal in a way because I think if you don't, people will just not use you know they will get fed up, stop doing it.* [LE29; male; aged 73 years]

Moreover, excessive interaction may also result in increased anxiety among patients, as constantly monitoring and checking symptoms may exacerbate their condition and worsen their quality of life:

> *If I keep constantly checking that machine, then I'm going to, and it's a little bit raised, or whatever I'm going to be continually worrying which doesn't help your blood pressure.* [LE5; female; aged 61 years]

Clinicians shared this concern, expressing reluctance to recommend a technology that could potentially cause further harm or anxiety for their patients:

> *It may backfire because the patient might get the wrong idea might get panic, might get anxious you know it might they might think they are getting feedback, it must be something very severe you know. So those things can be a backfire, you know they might get upset. They might get anxious.* [Clinician7; male; aged 44 years]

Finally, there was a consistent message across both participant groups that digital technologies could never truly replace face-to-face human contact, and any attempts to do so will ultimately worsen the quality of care across the cardiovascular care pathway:

> *I don't think you know a human face and a human voice will ever beat, you know will be beaten in the future. So I think you know we've got a struggle to do that, anyway.* [LE8; male; aged 61 years]

> *During COVID we found this because we thought, can we make use of some of these things? But what a lot of the patients said was missing actually was...more direct contact.* [Clinician6; female; aged 49 years]

### *There Is Too Much Individual-Level Variation*

There was consistent acknowledgment of the challenges related to individual-level variation and how this would inevitably impact engagement with any digital health technology. It is clear that both patients and clinicians can have very different experiences, beliefs, and familiarity with digital technologies, and it is difficult to implement technologies that suit the needs of every potential user, especially given the variation across heart diseases.

One patient reflected on how their heart disease requires very different care compared to others, highlighting the challenges of implementing effective digital technology within different heart disease diagnosis pathways:

> *I'm not particularly into wearable devices, because I think that they're probably far more useful for people who've got electrical problems with their heart, whereas mine is a plumbing issue, always has been.* [LE10; male; aged 65 years]

Clinicians also described how the nature of individual differences in preferences can act as a barrier to engagement, as it is not possible to suit everyone's needs, especially when it comes to different demographic factors and previous experiences:

> *Some patients are going to be up for it, and they would love to have something on their phone and they like, you know, there are patients who really like to record data, and they will love it. They will get their phone, and they'll get an app, and it will be fine. There are some who would be fairly resistant to it.* [Clinician2; female; aged 38 years]

Furthermore, clinicians expressed concerns regarding the accessibility of potential technologies, as any technology is heavily dependent on patients' understanding of the device or program, which often varies but can be difficult to predict on a larger scale:

> *So you have an app that can help to monitor the condition but the patient couldn't use it couldn't put in the data, then there's no point using those apps isn't it?* [Clinician7; male; aged 44 years]

## Theme 4: Resource-Related Facilitators

### *It Needs to Be Cost-Effective*

Clinicians considered evidence for the cost-effectiveness of a novel technology to be a facilitator of effective implementation; however, this was also dependent on adequate resources to support implementation from the relevant health care service or trust. This highlights the importance of considering financial implications from the costs to the individual to the costs to the health care system:

> *If it was going to be cost-effective you know, I don't have any way of bringing in new technology the way my practice works currently, you know...but it needs to be some way of bringing staff in to help me do things like that.* [Clinician13; male; aged 49 years]

### *It Needs to Be Efficient*

A key driver for engagement for both patients and clinicians related to the additional efficiency that health technologies could provide during the diagnosis process, as this could address current issues that are contributing to inaccurate or delayed heart disease diagnoses:

> *If it took the place of a 24-hour blood pressure monitoring or 24-hour ECG or what's your average pulse over this time, then actually, that's quite useful, because it's kind of doing, taking away some of the work or putting the workload elsewhere. It's doing the work that's already being done.* [Clinician2; female; aged 38 years]

Patients also shared how increasing efficiency would improve the quality of their health care experience and therefore act as an important facilitator of their engagement with novel technologies:

> *The automation of the whole process is, would be a blessing for me.* [LE10; male; aged 65 years]

> *I suppose it could be, if it's all digital data coming into one source that could be much more efficient.* [LE28; female; aged 50 years]

### *It Would Help to Have Setup Support*

There was a shared sentiment between both patients and clinicians regarding the importance of having adequate setup support at the initial point of implementation of any digital technology. In particular, clinicians highlighted that as it is not feasible for them to provide this support due to current resource limitations, they would be comforted by the knowledge that there is an external body responsible for supporting patients to set up the technology, as well as providing adequate support in case of technological issues at any stage:

> *If there was like a support line, they could ring instead, then, you know, we could just direct, you know, and say, actually, that's fine, or you will be contacted by the you know, this company will help you go through the app, then that's fine, I suppose.* [Clinician2; female; aged 38 years]

Patients also reflected that adequate provision is needed to make people feel confident in engaging in any health technology

related to their heart condition, with suggestions that language used in the setup support is crucial in increasing engagement among users:

> *I think you need somebody that's gonna help you. You need very plain un-jargonistic instructions so that we can follow it* [LE18; female; aged 66 years]

## Theme 5: Technology-Related Facilitators

### *Is it Going to Be Accurate and Reliable?*

Unsurprisingly, accuracy and reliability of technology were consistently brought up as important facilitators of engagement, as this elicits confidence in both clinicians and patients that they can use the technology to improve the quality of their experience or the accuracy of the diagnosis. Clinicians often expressed accuracy as the first thing they would consider when deciding whether to engage with a novel technology:

> *It should be accurate, I guess, accuracy is most important...good accuracy that would be ideal isn't it? So most of the data can be interpreted by a machine* [Clinician7; male; aged 44 years]

This was consistently echoed by patients, who felt accuracy was the foundation of a good digital health solution and would only agree to use something they were confident would produce accurate data that could be used within their health care pathway:

> *It would need to be very accurate.* [LE22; female; aged 68 years]

> *It's really hard to sort of summarize if you're having seen a clinician...you need to summarize quite a few weeks worth of data...[technology] is far more accurate trying to get a snapshot from a from any from a patient about their overall health, and especially their mental health.* [LE28; female; aged 50 years]

### *Safety Has Been Adequately Considered*

As mentioned previously, safety was a key area of discussion given the potential risks of monitoring symptoms before receiving a diagnosis. In fact, clinicians provided specific requirements for the way that data should be dealt with and thresholds that would need to be in place for them to feel confident in implementing novel technologies to aid in the diagnosis of heart diseases:

> *If it was kind of then inputting symptoms, it would have to have very strict criteria as to how it dealt with that. Yeah, I think, is the problem if it was just a manual thing that flashed up every time they entered, I have chest pain, you're going to have to be very careful what it said or did.* [Clinician2; female; aged 38 years]

Moreover, patients also shared their perspective on how data should be shared safely among the device, the patient, and the clinician, highlighting the nuance in the communication of risk and potentially concerning health data collected by a digital device:

> *Anything which goes above a certain level of importance, it should go to the doctors or medics or emergency services as required, but it has to be quite, shall we say a severe level to actually get to the giving out that warning.* [LE11; male; aged 70 years]

### *Is it Easy for me to Use?*

The consensus was that for any technology to be effectively implemented into clinical practice, it needs to be as simple as possible, as this produces the greatest level of widespread engagement and fewer complications for clinicians who need to use the data output:

> *Something that's easy to use...convenient to use, you know, for everybody, for the patient and us. Because then I know that they're more likely to use it.* [Clinician6; female; aged 49 years]

Patients also emphasized the importance of simplicity in novel technologies as well as making it easy to integrate them into current health care systems to ensure sustained engagement:

> *The key to get people to use anything is to make it easy. So, if we go down this route, which I think is great, we should piggy backing in on existing technologies...that can be used by every part of the NHS.* [LE17; male; aged 65 years]

### *Patients Have a Right to Their Data*

There was considerable discussion surrounding who should have access to health data collected by digital devices aiding in the diagnosis pathway; however, general attitudes of participants suggested that patients have a right to their own data, regardless of what they are being monitored for, as this encourages trust between the patient and the clinician:

> *I mean yeah it should be sent to patients and I think lots of, because that's the patient's information at the end of the day, and I guess a lot about health care is being open and transparent and actually you shouldn't be sending data out about a patient to the doctor and the patient not having that information.* [Clinician2; female; aged 38 years]

Interestingly, patients mainly expressed wanting clinicians to have access to their data, suggesting they did not feel confident in how to handle receiving their own health data without the support of a health care professional. This echoes previous concerns regarding safety and highlights the importance of making patients feel supported while depending on technology to collect and interpret their health data:

> *I would think the GP would be the first person to receive information and followed by myself and any associated to the medical profession, professional and in terms of when you refer to someone, a specialist, for example, if they're already involved. So that's the order that I would like to see it in.* [LE19; male; aged 64 years]

### *Personalization Is Key*

When considering the development of health technologies, personalization was a key element mentioned as a facilitator of

effective implementation. The clinicians' shared perspective highlighted the importance of making people feel that the technology was tailored toward them, instead of expecting people to tailor themselves to the technology. There was also a sense that past experiences had led to high expectations of technology, placing greater pressure on developers to design health technologies that align with public perceptions:

> *But yeah, generally speaking, people like stuff that they feel isn't just generic and sent out to everyone.* [Clinician2; female; aged 38 years]

Meanwhile, patients also emphasized the importance of receiving personalized and relevant data instead of generic feedback as a way of keeping people engaged. Patient perspectives also highlighted interest in examining trends and patterns within their health data, suggesting technologies should be designed based on the assumption that some patients may want to engage with their data beyond their clinical consultations:

> *What you'd want to do is to be able to interrogate the database that maybe there's some graphs and trends to see. You know how your reading is compared to average.* [LE10; male; aged 65 years]

## Theme 6: User-Related Facilitators

### *Adapting to Individual Characteristics*

Despite acknowledging how difficult it can be to develop health technologies tailored to individual differences, both patients and clinicians provided useful insights into how this could be done effectively to improve engagement. Clinicians emphasized the importance of asking patients how they wanted to interact with a digital technology as part of their diagnosis journey, as well as capturing clear expectations regarding their understanding and capabilities in relation to the technology as early as possible:

> *One way of addressing it is to ask the patient how much they would expect to interact. You know. That's one way to it, you know to ask the patient.* [Clinician7; male; aged 44 years]

> *I think the patients understanding the technology and being able to use it and to use it appropriately.* [Clinician9; male; aged 35 years]

Meanwhile, patients reflected on the importance of considering the target demographic when designing any health technology, as well as increased difficulties resulting from comorbidities:

> *But let's make it one device. So I don't have to have all the other devices. Otherwise they're going to be competing for my attention...I'm getting older and the target audience for this, most people who are ill are older, with multiple conditions.* [LE17; male; aged 65 years]

Overall, there was a clear message among participants that considering individual differences between patients is key to effective implementation and sustained engagement with novel health technologies aiming to improve heart disease diagnosis:

> *It also has to be, shall we say selective in what a single person or what the user requires it to do...so it has to be targeted individually to each individual person* [LE11; male; aged 70 years]

### *The Role of Positive User Attitudes*

It seemed that individual attitudes toward technology more generally, as well as its use in health care, played an influential role in willingness to engage with novel health technologies. Both patients and clinicians expressed a very positive outlook on the value of incorporating technologies into heart disease diagnosis, which translated as a greater willingness to engage:

> *I think, to be honest, the NHS, we need to go more and more towards these apps* [Clinician2; female; aged 38 years]

A crucial facilitator was also a perception that the technology would in fact be useful for them, whether this was based on evidence to show it would improve an aspect of their care or if they judged it as being a helpful addition based on past experiences:

> *It needs to be proved. It needs to be shown to some degree that it's definitely, it's making, improving the outcome before I use it.* [Clinician7; male; aged 44 years]

> *Yeah, I think that'd be good to have like a chat bot, where if you've got any questions or anything like that, you can just click and get them answered rather than having to try and wait and get in to see the doctor or a consultant.* [LE20; female; aged 54 years]

However, there was still a recurring sentiment that complete dependence on technology is not feasible, with patients emphasizing the importance of human oversight even if data are being collected remotely. This highlights a key aspect of digitalized health care that is important to stakeholders and should be considered thoroughly during implementation to increase engagement and create a sense of safety among participants:

> *I think what should happen is that the medical profession should be getting the feedback and react accordingly to that.* [LE29; male; aged 73 years]

### *It Improves the Quality of Patient Care*

Unsurprisingly, when stakeholders felt that they would experience direct benefits to the quality of their or their patients' care, they felt more motivated to engage with novel technologies. There were specific benefits that were mentioned by participants, with some degree of variation between patients and clinicians. Patients reflected on past experiences with health technologies, which made their lives easier because it made handling health data more convenient:

> *Any digital technology is advantageous both to the user and supplier. And I'll cite the Covid app, instead of carrying sheets and sheets of paper about with you if you go on holiday, on your Covid app, it tells you when you had it, where you had it, what it was that you got.* [LE1; male; aged 72 years]

Meanwhile, clinicians emphasized how having better access to their patients' health data made their jobs easier and allowed for better quality of care that was adapted to both clinician and patient needs:

> *I can access patients' information easier you know I don't have to be in the on the ward. It's just physically looking on the note, so it's a lot of, improves the flexibility.* [Clinician7; male; aged 44 years]

An improved access to health data also reduced anxiety in patients, as they expressed a feeling of relief for themselves and their families because of feeling more informed about their condition or their symptoms:

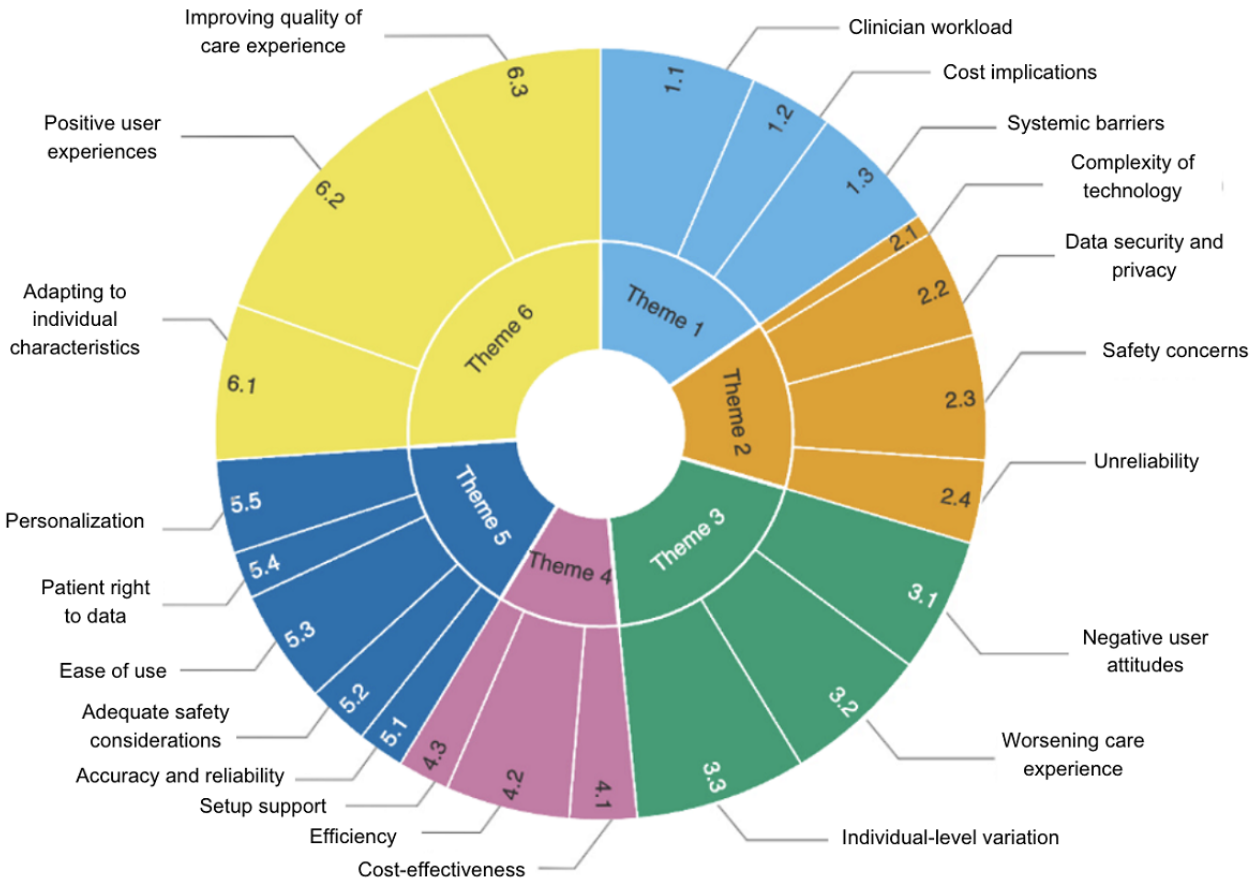
> *It just gives you peace of mind. And obviously with your family members. They put the knowledge around them as well...So that's it's a no brainer really. It's got to help.* [LE8; male; aged 61 years]

There was also evidence for a strong desire to be more involved in their own care pathway, as they felt this would improve their health care experiences and result in more transparency between the patient and the health care provider:

> *I would certainly welcome having more access to my medical records, because obviously, whenever I go and see a GP, I'm just amazed about how much data they've got about me, but I can't see it. I wish I could.* [LE10; male; aged 65 years]

Figure 2 presents the themes and subthemes described earlier in a sunburst diagram to illustrate the relative size of each subtheme within each of the 6 themes. This figure reveals that user-related barriers and facilitators (themes 3 and 6) emerged as the biggest themes, while resource-related barriers and facilitators (themes 1 and 4) were the smallest themes overall. Thus, these findings provide crucial insight to inform the development of novel health care technologies, particularly for the sake of making appropriate decisions to ensure user needs are met.

**Figure 2.** Sunburst visual of themes by size based on items coded, separated by themes, representing the barriers and facilitators of engagement with digital technologies for heart disease diagnosis.




## Recommendations

On the basis of the emerging themes presented earlier, we have developed recommendations that should be considered when developing digital technologies to assist in the diagnosis of cardiovascular diseases. These recommendations are divided into technology-specific considerations (related to how the technologies function or are used) and system-level considerations (how the broader health care system should adapt to successfully implement such technologies). Multimedia Appendix 4 summarizes these recommendations based on each theme that came from the data, collected from participants with

an interest in participating in digital technology research and clarified with support from the Patient Advisory Board.

## Discussion

### Principal Findings

This study has revealed the variety of barriers and facilitators influencing the effective implementation of digital technologies into the heart disease diagnosis pathway, as seen from the perspective of stakeholders with an interest in digital technology research. Both barriers and facilitators were organized into resource-, technology-, and user-related themes, with several subthemes within each of the 6 major themes.

Resource-related barriers and facilitators related to clinician workload, system-level influences, cost implications, efficiency, and support infrastructure. These findings are consistent with previous studies that have found increased clinician workload and a lack of integration into clinical workflow to be common barriers to the uptake of digital health technologies into cardiovascular care, while improved efficiency, institutional approval, and organizational support are all common facilitators [24,32]. Furthermore, technology-related barriers and facilitators included themes related to reliability, accuracy, safety parameters, data security, ease of use, and personalization. These perspectives were consistent with a recent qualitative review of wearable technology adoption for cardiac monitoring, which found 4 interrelated themes, including trust, safety and confidence, functionality and affordability, and risks and assurance [14]. Furthermore, concerns related to accessibility and usability of technology also emerged in a systematic review and content analysis of barriers and facilitators for health management across several physical and mental health conditions [33], highlighting the overlap in technology-related barriers among different stages of the care pathway. Overall, our findings emphasized key areas of technology development that could be adapted to improve the implementation of digital health technologies into the cardiovascular diagnosis pathway.

Finally, the most prominent themes were the user-related barriers and facilitators, which encompassed user attitudes, individual-level variations, and impact on quality of health care experiences. This theme captured a wide variety of perspectives among the sample and echoed findings from existing literature, which revealed how patient and clinician attitudes and personal experiences substantially impact engagement with digital health technologies across the cardiovascular care pathway, ranging from cardiac rehabilitation to remote care and self-management in heart failure [15,16,19,23]. These results also appear to be consistent across different clinical conditions, with a recent systematic review investigating barriers and facilitators to using digital health technologies finding that perceptions of usefulness and willingness to use novel technologies were important facilitators to enhance the uptake of digital health technologies by health care professionals across different clinical specialties [33]. Thus, the results of our study highlight the impact of user-related factors on the effective implementation of novel digital health technologies and therefore reveal a key area for future technology development to focus on to improve engagement levels during the diagnosis pathway.

Another key objective of this study was to understand potential differences between patient and clinician perspectives in relation to the barriers and facilitators mentioned earlier. Overall, the results of our study suggest that generally patients and clinicians share similar views on factors that may be preventing effective implementation of novel digital technologies into health care, as well as areas to focus on to facilitate better implementation. However, there were a few exceptions throughout the subthemes, with resource-related barriers (such as clinician workload and high costs) and technology-related safety concerns being discussed more by clinicians. Meanwhile, user-related barriers, such as negative attitudes toward technology and perceptions that quality of care would be reduced by novel technologies, were only presented as barriers by LE participants. These differences are consistent with the wider literature investigating factors influencing uptake of digital health technologies, as concerns related to resource restraints and evidence-based care also emerged as barriers in a sample of clinicians working with chronic obstructive pulmonary disease [34,35]. Moreover, while facilitators were mostly similar between both participant groups, the only exceptions were resource-related cost benefits and technology-related accuracy and reliability, which were facilitators emphasized by clinicians.

It is not surprising that clinicians presented more resource- and technology-related perspectives given they are more likely to be exposed to these aspects of novel technologies compared to patients [36]. It is also expected that patient perspectives would focus more on user experience and impact on quality of care, as they are able to draw on personal LE of how digital technologies used in their own care impacted their experiences. This distinction is consistent with the review by Whitelaw et al [24], which found that increased workload and a lack of integration with electronic medical records were identified as clinician-level barriers, while organizational support and improving efficiency were important facilitators according to clinician perspectives. A scoping review [32] focusing on hypertension management also found that concerns with integration of technologies into existing clinical workflow only emerged among health care professionals, while interference with patient- health care provider relationships was primarily a patient concern. Ultimately, our data highlight how different user groups may vary in which barriers are more influential in preventing them from engaging with health technologies within the heart disease diagnosis pathway. Therefore, the findings of this study provide useful insights into how implementation processes can be tailored to target these specific barriers, as well as consider facilitators, to increase uptake of novel health technologies within the heart disease diagnosis pathway.

The recommendations based on our qualitative findings for implementing health care technologies focused on addressing resource, technology, and user-related factors. Key strategies include integrating intuitive interfaces with existing IT systems, providing comprehensive training and support, and ensuring cost-effective models. Addressing technology-related barriers involves designing user-friendly, secure, and reliable systems with rigorous clinical trials and active monitoring for issues. Simplifying complexity and ensuring transparent data use are also essential. Facilitators for successful implementation include

demonstrating cost-effectiveness, improving efficiency, and offering extensive setup support for patients and clinicians. Ensuring accuracy and reliability through rigorous validation and regulatory frameworks, alongside enabling patient access to their data, is vital. Emphasizing personalization and adapting to individual user characteristics will further enhance user acceptance and improve the overall care experience. These considerations echo existing calls to address key issues associated with implementing technologies into clinical care, such as ensuring patients can trust the systems managing their data and clinicians are not overwhelmed by the large volume of data that are generated by wearable digital health technologies [37]. However, while these general recommendations provide a foundation, they may lack specificity when applied to certain contexts. For example, the type of heart diseases targeted by a digital diagnostic tool will influence not only its design but also its adoption and integration into existing care pathways. Similarly, the demographic and clinical characteristics of patients using the device, such as age, literacy, and comorbidities, may present unique challenges that require tailored solutions [38]. Finally, while the focus on cost-effectiveness and efficiency is commendable, these factors must be balanced against equity considerations. For example, ensuring access to these technologies for underserved populations or regions with limited resources is critical to avoid widening existing health care disparities. Therefore, a nuanced approach that considers these broader contextual, systemic, and equity-focused challenges is essential for the successful implementation of health care technologies [39].

### Strengths and Limitations

A key strength of this study was the use of a qualitative study design to capture both patient and clinician experiences. This depth of insight would not have been possible to achieve using quantitative methods. The use of a decentralized recruitment strategy for both participant groups also meant our sample included people from across the country and captured a range of health care and technology experiences. Moreover, patient and public involvement was intentionally incorporated into each stage of the study, from the creation of study materials to the review of preliminary thematic analysis results. This increases confidence that the study's design effectively created a comfortable environment for participants to share their experiences and ensured their data were interpreted accurately. While it is not possible to remove subjective bias from the lead researcher's interpretation and analysis of the qualitative data, the involvement of patient panels and LE advisers throughout the study can provide reassurance that the results are translatable beyond our sample.

However, there are several limitations that also need to be acknowledged. The web-based nature of our recruitment method may have resulted in a biased sample of individuals who were more confident using technology, meaning their experiences are unlikely to capture the challenges faced by patients and clinicians who have less experience with technologies. Moreover, we were not successful in recruiting *difficult to reach groups*, such as ethnic minority groups with different cultural experiences across the United Kingdom, despite efforts to use the research team's personal connections to include participants from underserved communities. This would have been extremely valuable to aid in our understanding of challenges related to accessibility and implementation of novel health technologies, so we suggest future research studies attempt to build on our findings and explore perspectives on barriers and facilitators in populations that are more resistant, or less experienced, in using digital health technologies. Our exclusion of people who were not fluent in English means our results exclude perspectives from people who may face different challenges and benefits from interacting with technology. An additional consideration is the differing forms of data collection. We made the pragmatic decision to run focus groups with LE participants and individual interviews with clinicians, due to the difficulties in getting multiple clinicians to be free at the same time for a focus group. This difference in data collection methods may have influenced results. Focus groups can result in more dynamic exchanges and can help foster a shared understanding of a phenomenon, resulting in different information shared than would be in an individual scenario. In contrast, interviews can allow for deeper, more personal insights to be shared [40]. While there is some precedent for the combination of qualitative methods, with researchers suggesting that it can be a useful method of triangulation to enhance depth and breadth of insights [41], there is ongoing debate about how different data collection methods can be most meaningfully combined in analysis. While we attempted to address this with our reflective approach to analysis, it is possible that our results and key findings may have differed if the same qualitative methods had been used to collect data from both LE and clinician participants.

Although the questions asked in focus groups and interviews were designed to be as vague and nonleading as possible, it should be acknowledged that this study was conducted as a part of a wider project aiming to develop a novel digital twin technology to improve holistic heart disease diagnosis. This meant the topic guides for both focus groups and interviews were focused on a specific technology being designed for a specific purpose; thus, it is possible that this may have excluded experiences and perspectives on other potential technologies that could be used within the heart disease diagnosis pathway.

Finally, we did not specifically recruit participants with direct experience of using digital technologies for health management. This intentional choice aimed to broaden the applicability of our findings; however, it may have impacted the nature of participants' responses, introducing a degree of hypothetical reasoning. However, even without direct experience of using these technologies or implementing them in health care services, all participants brought valuable insights based on their LEs with health care services, use of technologies in daily lives, and existing challenges in the system. Analytically, we handled this challenge by carefully interpreting the data within the scope of participants' experiences and triangulating results across multiple participants and sources to ensure that conclusions were not drawn from speculative responses.

## Conclusions

Digital technologies are a growing area, and our results provide insight into the key design and implementation characteristics needed to be accepted by patients and clinicians into routine

clinical care. This qualitative study has revealed the multifaceted barriers and facilitators influencing the implementation of digital technologies in the heart disease diagnosis pathway. The findings demonstrate that resource-, technology-, and user-related factors play critical roles in adoption, with user-related aspects emerging as particularly important. While patients and clinicians generally share similar perspectives on implementation challenges and opportunities, notable differences exist in their prioritization of specific barriers and facilitators. These insights emphasize the importance of tailored implementation strategies that address the unique concerns of both user groups. To increase the acceptability of novel health technologies in heart disease diagnosis, future developments should prioritize creating user-friendly, secure, and reliable systems that can be integrated into existing clinical infrastructure, as well as allowing for personalization and adaptability to individual user needs. Addressing these factors is key to fostering confidence in and uptake of digital diagnostic tools in cardiovascular care.

## Acknowledgments

The authors would like to thank the 2 patient and public involvement groups that helped to inform the design of this study: the National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre's Race, Ethnicity, and Diversity (READ) advisory group and the Sheffield-based Cardiology Patient group. The authors would also like to thank Helen Denney and Amber Ford for convening the Sheffield patient group and for administrative assistance and Manuel Cabeleria for helping to write the code for Figure 2. This work is supported by the UK Engineering and Physical Sciences Research Council (EP/X000257/1).

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

FM, TJC, JC, OB, VD, and RVA were responsible for the conceptualization. FM and TJC were responsible for methodology. KA, MM, and JC were responsible for the investigation. KA was responsible for writing the original draft. FM, TJC, JC, MM, JC, OB, VD, and RVA were responsible for writing—review and editing. FM and TJC were responsible for supervision. KA and FM were responsible for project administration. FM, TJC, JC, OB, VD, and RVA were responsible for acquisition.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Question schedule.
[PDF File (Adobe PDF File), 92 KB - cardio_v9i1e66464_app1.pdf ]

Multimedia Appendix 2
Information sheets and consent forms.
[PDF File (Adobe PDF File), 370 KB - cardio_v9i1e66464_app2.pdf ]

Multimedia Appendix 3
Participant demographics questionnaire.
[PDF File (Adobe PDF File), 100 KB - cardio_v9i1e66464_app3.pdf ]

Multimedia Appendix 4
Themes and subthemes.
[DOCX File , 18 KB - cardio_v9i1e66464_app4.docx ]

## References

1. Bhatia RT, Gati S, Papadakis M, Sharma S. The impact of COVID-19 on the continuity of cardiovascular care. Eur Heart J 2021 Jan 20;42(3):215-217 [FREE Full text] [doi: 10.1093/eurheartj/ehaa742] [Medline: 33021673]
2. Harky A, Adan A, Mohamed M, Elmi A, Theologou T. Technology and cardiovascular diseases in the era of COVID-19. J Card Surg 2020 Dec 10;35(12):3551-3554 [FREE Full text] [doi: 10.1111/jocs.15096] [Medline: 33040461]
3. Bouabida K, Lebouché B, Pomey MP. Telehealth and COVID-19 pandemic: an overview of the telehealth use, advantages, challenges, and opportunities during COVID-19 pandemic. Healthcare (Basel) 2022 Nov 16;10(11):2293 [FREE Full text] [doi: 10.3390/healthcare10112293] [Medline: 36421617]

4. Chico TJ, Stamatakis E, Ciravegna F, Dunn J, Redwood S, Al-Lamee R, et al. Device-based measurement of physical activity in cardiovascular healthcare: possibilities and challenges. Br J Sports Med 2023 Oct 07;57(19):1225-1226 [FREE Full text] [doi: 10.1136/bjsports-2022-106460] [Medline: 37549998]
5. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019 Nov 14;381(20):1909-1917. [doi: 10.1056/nejmoa1901183]
6. Orini M, van Duijvenboden S, Young WJ, Ramírez J, Jones AR, Hughes AD, et al. Long-term association of ultra-short heart rate variability with cardiovascular events. Sci Rep 2023 Nov 03;13(1):52. [doi: 10.1038/S41598-023-45988-2]
7. Kędzierski K, Radziejewska J, Sławuta A, Wawrzyńska M, Arkowski J. Telemedicine in cardiology: modern technologies to improve cardiovascular patients' outcomes-a narrative review. Medicina (Kaunas) 2022 Feb 01;58(2):210 [FREE Full text] [doi: 10.3390/medicina58020210] [Medline: 35208535]
8. Stremmel C, Breitschwerdt R. Digital transformation in the diagnostics and therapy of cardiovascular diseases: comprehensive literature review. JMIR Cardio 2023 Aug 30;7:e44983 [FREE Full text] [doi: 10.2196/44983] [Medline: 37647103]
9. Coorey G, Figtree GA, Fletcher DF, Snelson VJ, Vernon ST, Winlaw D, et al. The health digital twin to tackle cardiovascular disease-a review of an emerging interdisciplinary field. NPJ Digit Med 2022 Aug 26;5(1):126 [FREE Full text] [doi: 10.1038/s41746-022-00640-7] [Medline: 36028526]
10. Bruynseels K, de Sio FS, van den Hoven J. Digital twins in health care: ethical implications of an emerging engineering paradigm. Front Genet 2018 Feb 13;9:31 [FREE Full text] [doi: 10.3389/fgene.2018.00031] [Medline: 29487613]
11. Elayan H, Aloqaily M, Guizani M. Digital twin for intelligent context-aware IoT healthcare systems. IEEE Internet Things J 2021 Dec 1;8(23):16749-16757. [doi: 10.1109/jiot.2021.3051158]
12. Abdullayev K, Gorvett O, Sochiera A, Laidlaw L, Chico T, Manktelow M, et al. Stakeholder perspectives on contributors to delayed and inaccurate diagnosis of cardiovascular disease and their implications for digital health technologies: a UK-based qualitative study. BMJ Open 2024 May 20;14(5):e080445 [FREE Full text] [doi: 10.1136/bmjopen-2023-080445] [Medline: 38772579]
13. Jiang X, Ming WK, You JH. The cost-effectiveness of digital health interventions on the management of cardiovascular diseases: systematic review. J Med Internet Res 2019 Jun 17;21(6):e13166 [FREE Full text] [doi: 10.2196/13166] [Medline: 31210136]
14. Ferguson C, Hickman LD, Turkmani S, Breen P, Gargiulo G, Inglis SC. "Wearables only work on patients that wear them": barriers and facilitators to the adoption of wearable cardiac monitoring technologies. Cardiovasc Digit Health J 2021 Apr;2(2):137-147 [FREE Full text] [doi: 10.1016/j.cvdhj.2021.02.001] [Medline: 35265900]
15. Tadas S, Coyle D. Barriers to and facilitators of technology in cardiac rehabilitation and self-management: systematic qualitative grounded theory review. J Med Internet Res 2020 Nov 11;22(11):e18025 [FREE Full text] [doi: 10.2196/18025] [Medline: 33174847]
16. Anttila MR, Kivistö H, Piirainen A, Kokko K, Malinen A, Pekkonen M, et al. Cardiac rehabilitees' technology experiences before remote rehabilitation: qualitative study using a grounded theory approach. J Med Internet Res 2019 Feb 07;21(2):e10985 [FREE Full text] [doi: 10.2196/10985] [Medline: 30730298]
17. Sivakumar B, Lemonde M, Stein M, Mak S, Al-Hesayen A, Arcand J. Patient perspectives on the use of mobile apps to support heart failure management: a qualitative descriptive study. PLoS ONE 2023 May 11;18(5):e0285659. [doi: 10.1371/journal.pone.0285659]
18. Wali S, Keshavjee K, Nguyen L, Mbuagbaw L, Demers C. Correction: using an electronic app to promote home-based self-care in older patients with heart failure: qualitative study on patient and informal caregiver challenges. JMIR Cardio 2020 Nov 11;4(1):e25624 [FREE Full text] [doi: 10.2196/25624] [Medline: 33175695]
19. Giordan LB, Ronto R, Chau J, Chow C, Laranjo L. Use of mobile apps in heart failure self-management: qualitative study exploring the patient and primary care clinician perspective. JMIR Cardio 2022 Apr 20;6(1):e33992 [FREE Full text] [doi: 10.2196/33992] [Medline: 35442205]
20. Eslick GD, Jones MP, Talley NJ. Non‐cardiac chest pain: prevalence, risk factors, impact and consulting — a population‐based study. Aliment Pharmacol Ther 2003 May 02;17(9):1115-1124. [doi: 10.1046/j.1365-2036.2003.01557.x]
21. Stull DE, Starling R, Haas G, Young JB. Becoming a patient with heart failure. Heart Lung 1999 Jul;28(4):284-292. [doi: 10.1016/s0147-9563(99)70075-0] [Medline: 10409315]
22. McSweeney JC, Lefler LL, Crowder BF. What's wrong with me? Women's coronary heart disease diagnostic experiences. Prog Cardiovasc Nurs 2005;20(2):48-57. [doi: 10.1111/j.0889-7204.2005.04447.x] [Medline: 15886547]
23. Al-Naher A, Downing J, Scott KA, Pirmohamed M. Factors affecting patient and physician engagement in remote health care for heart failure: systematic review. JMIR Cardio 2022 Apr 06;6(1):e33366 [FREE Full text] [doi: 10.2196/33366] [Medline: 35384851]
24. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, van Spall HG. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74 [FREE Full text] [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]
25. Abdullayev K, Chico TJ, Manktelow M, Buckley O, Condell J, van Arkel RJ, et al. Stakeholder-led understanding of the implementation of digital technologies within heart disease diagnosis: a qualitative study protocol. BMJ Open 2023 Jun 27;13(6):e072952 [FREE Full text] [doi: 10.1136/bmjopen-2023-072952] [Medline: 37369399]

26. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
27. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):25. [doi: 10.1186/S12874-018-0594-7]
28. Hennink M, Kaiser BN. Sample size for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523. [doi: 10.1016/j.socscimed.2021.114523]
29. Matcham F, Barattieri di San Pietro C, Bulgari V, de Girolamo G, Dobson R, Eriksson H, et al. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. BMC Psychiatry 2019 Feb 18;19(1):e514. [doi: 10.1186/s12888-019-2049-z]
30. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a researcher. Int J Transgend Health 2023 Oct 25;24(1):1-6 [FREE Full text] [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
32. Palacholla RS, Fischer N, Coleman A, Agboola S, Kirley K, Felsted J, et al. Provider- and patient-related barriers to and facilitators of digital health technology adoption for hypertension management: scoping review. JMIR Cardio 2019 Mar 26;3(1):e11951 [FREE Full text] [doi: 10.2196/11951] [Medline: 31758771]
33. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
34. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, Martinez EZ, Zucoloto ML, Østengaard L, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):61. [doi: 10.1038/s41746-023-00899-4]
35. Slevin P, Kessie T, Cullen J, Butler MW, Donnelly SC, Caulfield B. Exploring the barriers and facilitators for the use of digital health technologies for the management of COPD: a qualitative study of clinician perceptions. QJM 2020 Mar 01;113(3):163-172. [doi: 10.1093/qjmed/hcz241] [Medline: 31545374]
36. Yang WE, Shah LM, Spaulding EM, Wang J, Xun H, Weng D, et al. The role of a clinician amid the rise of mobile health technology. J Am Med Inform Assoc 2019 Nov 01;26(11):1385-1388 [FREE Full text] [doi: 10.1093/jamia/ocz131] [Medline: 31373364]
37. Ginsburg GS, Picard RW, Friend SH. Key issues as wearable digital health technologies enter clinical care. N Engl J Med 2024 Mar 21;390(12):1118-1127. [doi: 10.1056/NEJMra2307160] [Medline: 38507754]
38. Knisely BM, Vaughn-Cooke M, Wagner LA, Fink JC. Device personalization for heterogeneous populations: leveraging physician expertise and national population data to identify medical device patient user groups. User Model User-Adap Inter 2021 Oct 12;31(5):979-1025. [doi: 10.1007/S11257-021-09305-8]
39. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144 [FREE Full text] [doi: 10.2196/34144] [Medline: 35311682]
40. Guest G, Namey E, Taylor J, Eley N, McKenna K. Comparing focus groups and individual interviews: findings from a randomized study. Int J Soc Res Methodol 2017 Feb 13;20(6):693-708. [doi: 10.1080/13645579.2017.1281601]
41. Lambert SD, Loiselle CG. Combining individual interviews and focus groups to enhance data richness. J Adv Nurs 2008 Apr 03;62(2):228-237. [doi: 10.1111/j.1365-2648.2007.04559.x]


## Abbreviations

**COREQ:** Consolidated Criteria for Reporting Qualitative Research
**ECG:** electrocardiogram
**LE:** lived experience
**NHS:** National Health Service

*Edited by A Coristine; submitted 13.09.24; peer-reviewed by EW Verkerk, JP Gavin, I Wilson; comments to author 06.12.24; revised version received 15.01.25; accepted 04.02.25; published 05.03.25.*

*Please cite as:*
*Abdullayev K, Chico TJA, Canson J, Mantelow M, Buckley O, Condell J, Van Arkel RJ, Diaz-Zuccarini V, Matcham F*
*Exploring Stakeholder Perspectives on the Barriers and Facilitators of Implementing Digital Technologies for Heart Disease Diagnosis: Qualitative Study*
*JMIR Cardio 2025;9:e66464*
*URL: https://cardio.jmir.org/2025/1/e66464*
*doi:10.2196/66464*
*PMID:40053721*

# Application of Dragonnet and Conformal Inference for Estimating Individualized Treatment Effects for Personalized Stroke Prevention: Retrospective Cohort Study

Sermkiat Lolak[1], MD, PhD; John Attia[2], MD, Prof Dr, PhD; Gareth J McKay[3], MD, PhD; Ammarin Thakkinstian[4], PhD, Prof Dr

[1]Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, 4th Floor, Sukho Place Building, 218/11 Sukhothai Road, Suan Chitlada, Dusit, Thailand

[2]Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, Hunter Medical Research Institute, University of Newcastle, New Lambton, New South Wales, Australia

[3]Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, United Kingdom

[4]Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

**Corresponding Author:**
Sermkiat Lolak, MD, PhD
Department of Clinical Epidemiology and Biostatistics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, 4th Floor, Sukho Place Building, 218/11 Sukhothai Road, Suan Chitlada, Dusit, Thailand

## Abstract

**Background:** Stroke is a major cause of death and disability worldwide. Identifying individuals who would benefit most from preventative interventions, such as antiplatelet therapy, is critical for personalized stroke prevention. However, traditional methods for estimating treatment effects often focus on the average effect across a population and do not account for individual variations in risk and treatment response.

**Objective:** This study aimed to estimate the individualized treatment effects (ITEs) for stroke prevention using a novel combination of Dragonnet, a causal neural network, and conformal inference. The study also aimed to determine and validate the causal effects of known stroke risk factors—hypertension (HT), diabetes mellitus (DM), dyslipidemia (DLP), and atrial fibrillation (AF)—using both a conventional causal model and machine learning models.

**Methods:** A retrospective cohort study was conducted using data from 275,247 high-risk patients treated at Ramathibodi Hospital, Thailand, between 2010 and 2020. Patients aged >18 years with HT, DM, DLP, or AF were eligible. The main outcome was ischemic or hemorrhagic stroke, identified using *International Classification of Diseases, 10th Revision* (*ICD-10*) codes. Causal effects of the risk factors were estimated using a range of methods, including: (1) propensity score–based methods, such as stratified propensity scores, inverse probability weighting, and doubly robust estimation; (2) structural causal models; (3) double machine learning; and (4) Dragonnet, a causal neural network, which was used together with weighted split-conformal quantile regression to estimate ITEs.

**Results:** AF, HT, and DM were identified as significant stroke risk factors. Average causal risk effect estimates for these risk factors ranged from 0.075 to 0.097 for AF, 0.017 to 0.025 for HT, and 0.006 to 0.010 for DM, depending on the method used. Dragonnet yielded causal risk ratios of 4.56 for AF, 2.44 for HT, and 1.41 for DM, which is comparable to other causal models and the standard epidemiological case-control study. Mean ITE analysis indicated that several patients with DM or DM with HT, who were not receiving antiplatelet treatment at the time of data collection, showed reductions in total risk of −0.015 and −0.016, respectively.

**Conclusions:** This study provides a comprehensive evaluation of stroke risk factors and demonstrates the feasibility of using Dragonnet and conformal inference to estimate ITEs of antiplatelet therapy for stroke prevention. The mean ITE analysis suggested that those with DM or DM with HT, who were not receiving antiplatelet treatment at the time of data collection, could potentially benefit from this therapy. The findings highlight the potential of these advanced techniques to inform personalized treatment strategies for stroke, enabling clinicians to identify individuals who are most likely to benefit from specific interventions.

***(JMIR Cardio 2025;9:e50627)*** doi:10.2196/50627

**KEYWORDS**

stroke; causal effect; ITE; individual treatment effect; Dragonnet; conformal inference; mortality; hospital records; hypertension; risk factor; diabetes; dyslipidemia; atrial fibrillation; machine learning; treatment

## Introduction

Stroke is a leading cause of death and disability, presenting both personal and economic burdens [1]. Astonishingly, many epidemiological studies have identified important risk factors of stroke occurrence, especially through the use of cohort studies [2], and randomized controlled trials (RCTs) have identified the impact of treating these risk factors. While RCTs control for confounding factors through study design, cohort studies attempt to address these factors using statistical methods. However, the possibility of residual confounding remains, highlighting the need for improved analysis approaches [3].

Frameworks of causal effect have largely been confined to Pearl's [4] structural causal models (SCMs) and Rubin's [5] potential outcome models (POMs) [6]. SCMs evaluate causal relationships between variables using a directed acyclic graph defined by a set of structural equations, which consider the influence of each variable by its parents, or causes, along with its probability distribution. In addition, SCMs can also assess the effect of interventions by estimating how changing one unit of treatment (or risk) leads to a change in outcome [7]. Conversely, POMs focus on the concept of counterfactuals, specifically what would have happened if an individual had been exposed to a different treatment or risk [8]. Consequently, this approach estimates 2 potential outcomes (POs) for each individual: if the individual had received the treatment and if they had not. Subsequently, Rosenbaum and Rubin [9] developed propensity scores to reflect the probability of an individual being assigned to a certain treatment group. Therefore, these estimates are only considered valid if the 2 specific conditions—strong ignorability and positivity—are met. Statistical methods have been developed based on POMs and propensity scores, including matching [10], stratified propensity score (SPS) [11], inverse probability weighting (IPW) [12,13], and doubly robust estimation (DRE) [14-16]. Recently, nonconventional statistical models such as double machine learning (DML), meta-learners, and neural networks have also been developed to estimate unbiased causal effects without requiring strong underlying assumptions [14]. Causal neural networks (NNs), including TARNet and Dragonnet, learn by sharing input data to estimate both factual and counterfactual outcomes. This approach is currently an active area of research [17-19]. Dragonnet also uses "learned data" to predict propensity scores by tradeoff with prediction quality, which yields better average treatment effect (ATE) estimates [18].

Current causal modeling has shifted its focus from the ATE, which measures the treatment effect averaged across the entire study population, to the conditional average treatment effect (CATE), which assesses the ATE conditional on particular variables, such as sex, age, and other covariates. More recently, the focus has further evolved to the individualized treatment effect (ITE), which estimates the treatment effect for a particular individual. CATE has inherent variability depending on which covariate the model is conditioned on [20]. However, estimating ITEs is challenging because it requires making assumptions about the underlying individual data-generating process and the model used to estimate the ITEs [17]. A statistical technique called conformal inference may appropriately estimate the confidence intervals of ITEs by accounting for the uncertainty in their estimation. Despite being a novel technique, it has shown promise [20]. Conformal inference uses nonconformity scores that measure the degree of disagreement between the estimated and observed outcomes, to provide a confidence interval or a precision of estimation [21-23]. Therefore, we conducted this study to estimate the CATE of stroke occurrence based on real-world clinical data using Dragonnet NN models. Additionally, ITE was estimated to identify individuals at high risk of stroke who may benefit from lowering risk factors by combining the strengths of Dragonnet and conformal inference approaches. To the best of our knowledge, no prior studies have employed these methods in combination to estimate causal effects in a clinical setting.

## Methods

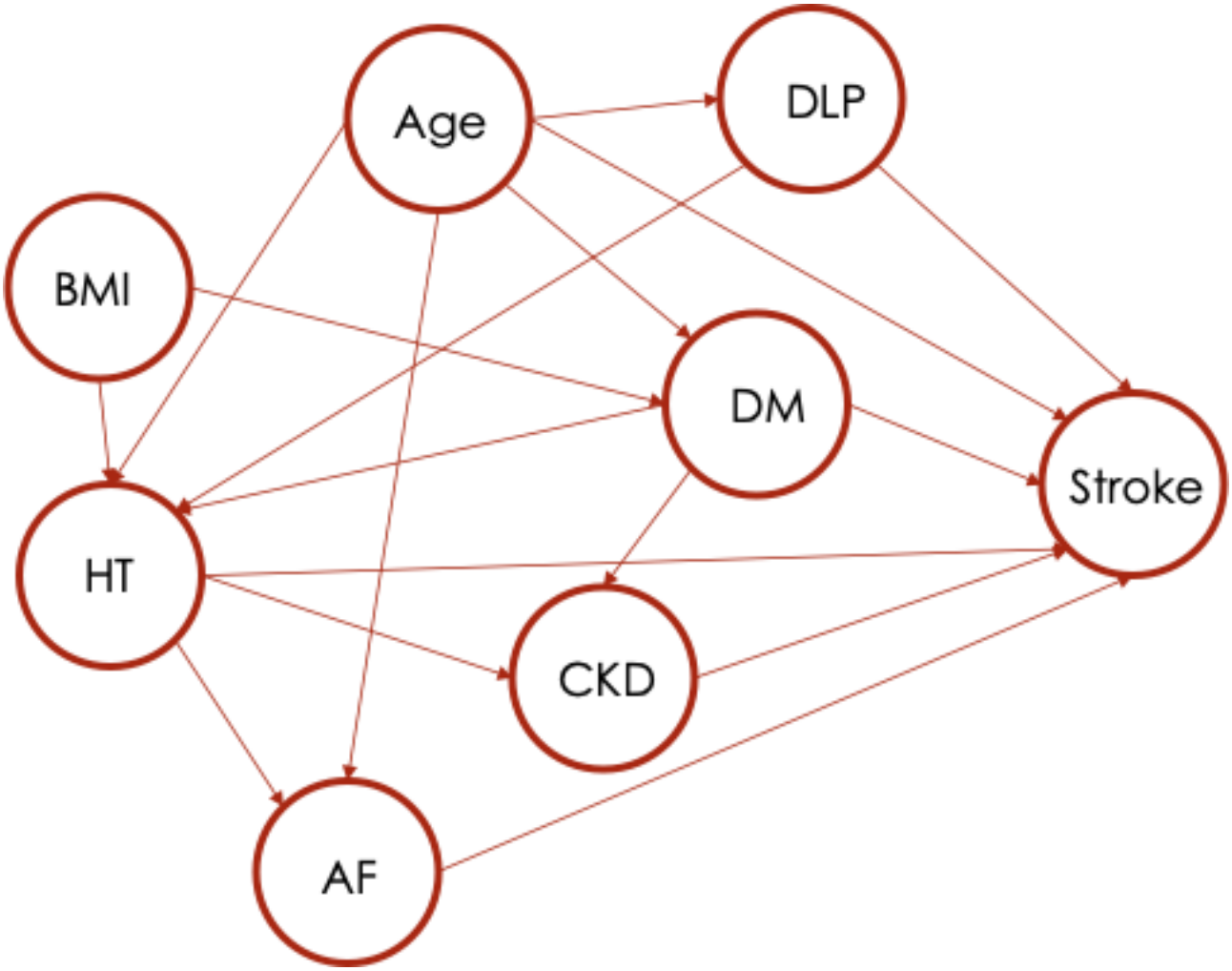
### Overview

The study population included a retrospective cohort of patients who were at high risk for stroke and had been treated and followed up at Ramathibodi Hospital, Thailand, between 2010 and 2020. Hospital records and the *International Classification of Diseases, 10th Revision* (*ICD-10*) classification system were used to identify patients. Patients were eligible if they were aged >18 years and had one or more of the following conditions: hypertension (HT; *ICD-10* code I10-I16), diabetes mellitus (DM; *ICD-10* code E08-E13), dyslipidemia (DLP; *ICD-10* code E78), and atrial fibrillation (AF; *ICD-10* code I48). Patients were excluded if they had a stroke on their first visit or only had one visit during the study period. The main outcome measured in the study was the occurrence of ischemic or hemorrhagic stroke, which was identified using the *ICD-10* codes I63 and I61, respectively.

Patients were followed up from their index date (i.e., the date they were identified as high-risk patients) until they progressed to stroke, were lost to follow-up, or were stroke-free at the end of the study (December 31, 2020). Patients who were lost to follow-up or stroke-free at the end of the study period were censored on their last visit date or at the end of the study. A causal diagram was constructed (Figure 1), and potential predictors of stroke were collected, including age, sex, BMI, chronic kidney disease (CKD), AF, HT, DM, and DLP. HT, AF, and DM were considered as mediators, whereas the remaining variables were covariates in the models. A software library called DoWhy, now incorporated into PyWhy (Python Software Foundation), was used to construct models for stratification, IPW, DRE, and DML [24]. Parameters of all estimators were set by default in the DoWhy package. The number of strata in the stratification method was automatically determined [25]. The weighting scheme in IPW was set to default inverse propensity score. For DRE, the regression and propensity models were specified as lasso and logistic regression, respectively. For DML, linear and nonlinear cross-fitted models were applied to the outcome model (lasso and Extreme Gradient Boosting [XGBoost]), propensity model (logistic regression and XGBoost), and final model (linear regression and lasso). Estimands of each risk pathway were defined by PyWhy from the input causal graph. Graphical causal

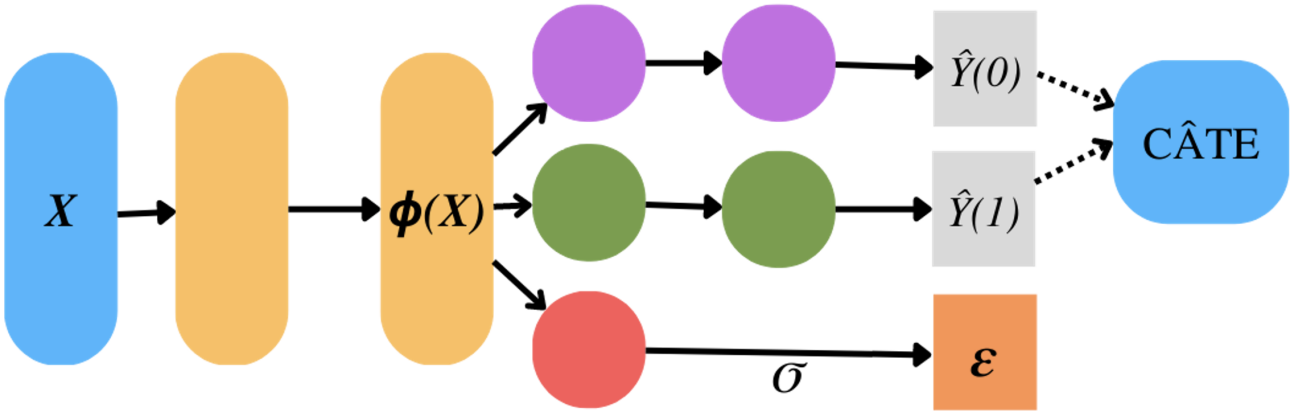
model–based inferences from the DoWhy library were used for medication analysis to quantify the causal effects of direct and indirect pathways, termed natural direct effect (NDE) and natural indirect effect (NIE), respectively [4,26]. NDE (Y1,M(0)x−Y0,M(0)x) refers to the change in the outcome of an individual when they are exposed to a specific treatment Y1, compared to another treatment Y0, while keeping the mediator variable constant at the baseline value or reference treatment *M(0)*. In contrast, NIE (Y1,M(1)x-Y1,M(0)x) refers to the difference between the counterfactual outcome value when treatment Y1 is fixed and the mediator assumes a certain value at a particular treatment *M(1)* and the counterfactual outcome value when the mediator assumes the same value at the baseline *M(0)* [27].

**Figure 1.** Causal diagram of patients at risk of stroke occurrence. AF: atrial fibrillation; CKD: chronic kidney disease; DLP: dyslipidemia; DM: diabetes mellitus; HT: hypertension.




The Dragonnet NN was used to estimate PO and propensity scores. The architecture of Dragonnet was based on previous work (Figure 2) [18]. It employs a deep net to create a representation layer *(X)* ∈ , which is used to forecast outcomes for both the treatment $\hat{Y}(1)$ and control groups $\hat{Y}(0)$. It utilizes 2 hidden layers for each outcome model while a basic fully connected layer with a sigmoid function is used for the propensity score ( ). CATE was estimated by subtracting treatment (risk) and control PO for each risk factor (Y1x−Y0x | Z) and risk ratios were estimated by division of PO (Y1xY0x|Z); $Y_1$ is the PO for the risk group, *Y* is the PO for the control group, *x* is an interested factor, and *Z* are other covariates.

**Figure 2.** Dragonnet architecture. *X* is the covariates, *(X)* is a learned representation of *X*. $\hat{Y}(1)$ is the predicted outcome of the treatment (risked) group. $\hat{Y}(0)$ is the predicted outcome of the control group. ε is the estimated propensity score. CÂTE is the conditional average treatment effect computed by $\hat{Y}(1)–\hat{Y}(0)$.




To accurately estimate the ITE, it is mandatory for the conditional independence assumption to hold, especially considering the unequal distribution of covariates between factual and counterfactual outcomes of the treatment and control groups, commonly known as covariate shift. To address this challenge, we employed a nested method of weighted split-conformal quantile regression (CQR) to estimate the ITE [20,23] by incorporating antiplatelet medications as a treatment for stroke prevention. POs were estimated using quantile loss setting α at .05. The dataset was split evenly into training and evaluation sets; Multimedia Appendix 1 shows the entire algorithm. All risk factors and covariates were similar between models, considering antiplatelet medication as a treatment and stratified by risk factor (Yantiplatelets=1x-Yantiplatelets=0x|Z), with *x* representing the risk factors of interest (i.e., HT, DM, and DLP) and Z representing other covariates. AF was not included as a stratum for the estimation of ITE in this example since it is not an indication for the prescription of antiplatelet therapy, but it remained a covariate.

### Ethical Considerations

The data were anonymized to ensure confidentiality and privacy protection. This study was approved by the Human Research Ethics Committee, Faculty of Medicine Ramathibodi Hospital, Mahidol University (COA. MURA2021/255). The committee waived the need to obtain consent for the collection, analysis, and publication of the retrospectively obtained and anonymized data for this noninterventional study.

## Results

A total of 275,247 high-risk patients were included in the cohort. Among them, 9659 patients developed stroke, resulting in an incidence of 3.5% (95% CI 3.4-3.6). The follow-up rate for the study population was 80% (7752/9659).

Baseline demographic and risk factors were compared between 9659 stroke patients and 265,588 nonstroke patients (Multimedia Appendix 2). Stroke patients had a mean age of 64.7 years and were more likely to be male. Stratification by risk indicated that 13% of AF patients, 4% of HT patients, 4% of DM patients, and 4% of DLP patients experienced stroke in contrast to only 2% of non-AF patients, 1% of non-HT patients, 3% of non-DM patients, and 3% of non-DLP patients, who developed stroke.

Causal effects of mediators including HT, DM, CKD, and AF on stroke were estimated based on the causal diagram in Figure 1. The estimands report as probability of stroke given the risk factors, *P*(*Stroke* | *risk factors*), are as follows: *P*(*Stroke* | *HT, age, DM, DLP*) for HT; *P*(*Stroke* | *AF, age, HT*) for AF; *P*(*Stroke* | *age, DLP*) for DLP; and *P*(*Stroke* | *age, DM, BMI*) for DM (Multimedia Appendix 3). For the POM approach, the SPS estimator showed AF as the highest risk of stroke, followed by HT, DM, and DLP with risk estimates of 0.084 (95% CI 0.079-0.088), 0.019 (95% CI 0.015-0.020), 0.010 (95% CI 0.008-0.010), and 0.0015 (95% CI −0.0002 to 0.0027), respectively. IPW yielded similar, albeit slightly higher, corresponding risks of 0.092 (95% CI 0.089-0.096), 0.024 (95% CI 0.022-0.025), 0.010 (95% CI 0.008-0.010), and 0.001 (95% CI −0.0005 to 0.0025), respectively. Comparable results were observed in the DRE analysis, with a similar trend of risk effect estimates of 0.082 (95% CI 0.0849-0.0871), 0.025 (95% CI 0.0243-0.0257), 0.008 (95% CI 0.0057-0.0063), and 0.0006 (95% CI 0.0001-0.0011), respectively.

The SCM estimation also yielded similar trends to the POM approach, in which the risk of stroke was 0.096 (95% CI 0.0948-0.0972), 0.021 (95% CI 0.0204-0.0216), 0.007 (95% CI 0.0067-0.0073), and 0.0005 (95% CI 0.0004-0.0006) for AF, HT, DM, and DLP, respectively. Mediation analysis indicated the NDE of HT to be 0.020 (95% CI 0.019-0.021) and the NIE to be 0.0027 (95% CI 0.0025-0.0029). NDE and NIE for DM and DLP were both modest and consistent with the findings from other models. Figure 1 illustrates the pathways through which the mediators act: HT mediates through CKD and AF, DM mediates through HT and CKD, while DLP mediates through HT.

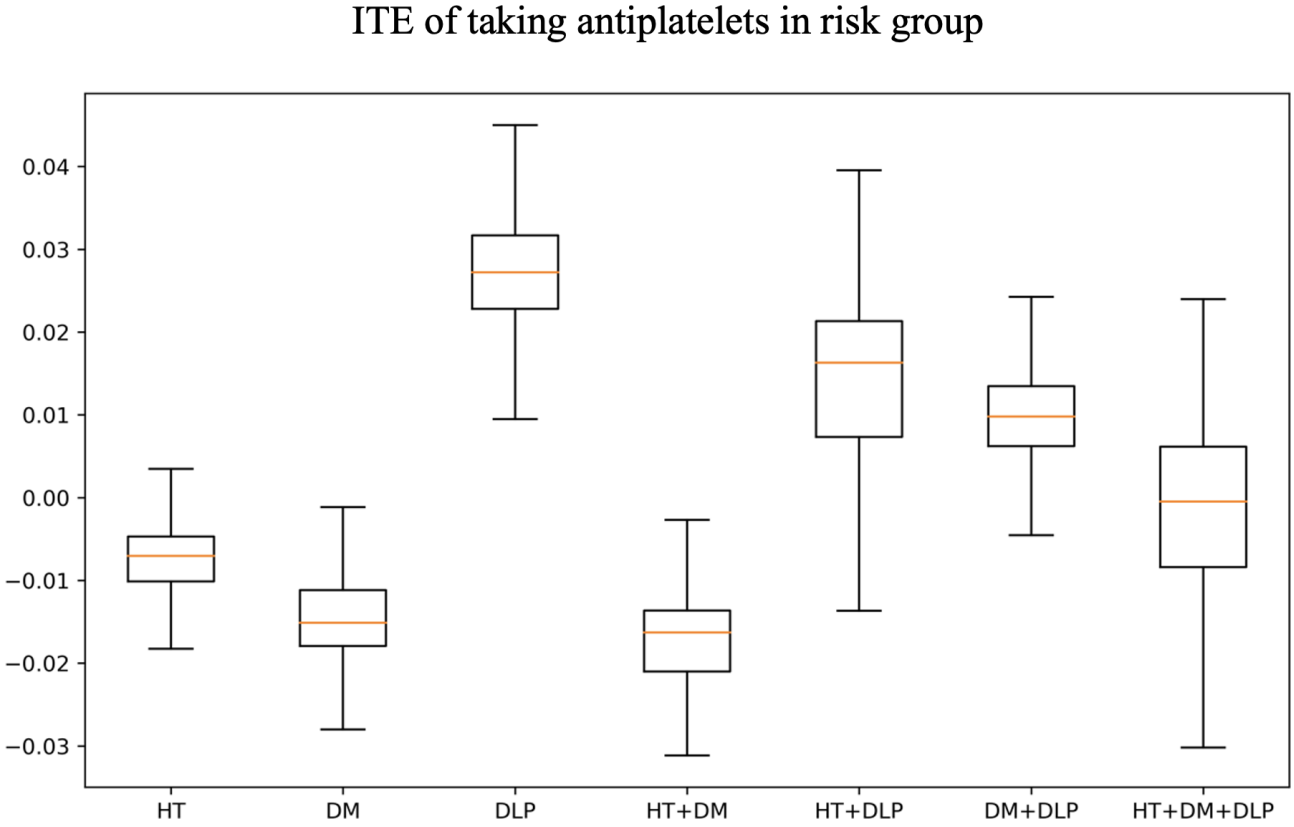
In the context of DML, the nonparametric model estimates were slightly smaller than those for the linear model, with risks of 0.086 (95% CI 0.0849-0.0871), 0.015 (95% CI 0.0145-0.0155), 0.006 (95% CI 0.0057-0.0063), and 0.0 (95% CI −0.0001 to 0.001) for AF, HT, DM, and DLP, respectively, whereas the corresponding linear model estimate risks were 0.097 (95% CI 0.096-0.098), 0.023 (95% CI 0.0223-0.0236), 0.009 (95% CI 0.0087-0.0093), and 0.002 (95% CI 0.0018-0.0022).

Dragonnet estimated the causal effects of AF, HT, DM, and DLP on stroke as 0.075 (95% CI 0.074-0.076), 0.017 (95% CI

0.0169-0.0170), 0.01 (95% CI 0.009-0.010), and −0.002 (95% CI −0.0022 to 0.0021), with causal ratios of 4.56 (95% CI 4.56-4.57), 2.44 (95% CI 2.41-2.46), 1.41 (95% CI 1.21-1.60), and 0.856 (95% CI 0.855-0.858), respectively. The odds ratios from the logistic regression models were respectively 3.34 (95% CI 2.68-3.75), 2.56 (95% CI 2.33-2.80), 1.16 (95% CI 1.05-1.30), and 1.00 (95% CI 0.8-1.4). Details are provided in Multimedia Appendix 3 for comparison.

The influence of risk reduction for individual patients who did not receive antiplatelet therapy, had they been given the medication (counterfactuals of nontreatment ITEs), was examined using weighted split-CQR. As shown in Multimedia Appendix 4, three of the samples (3/50, 6%) appear to have potentially benefited from antiplatelet treatment, indicating that a considerable number of patients might have experienced a positive impact on their stroke risk reduction had they received the medication. The mean ITEs indicated that several patients with DM or DM with HT were not currently receiving antiplatelet treatment and would be more likely to benefit if they had received it, with reduction of total risk as −0.015 (IQR −0.011 to −0.018) and −0.016 (IQR −0.015 to 0.022) among each group, respectively (Figure 3).

**Figure 3.** Box plot representing the mean individual treatment effect for patients with different risk factors who had not been taking antiplatelet medication, illustrating the potential impact on stroke risk reduction if they had received antiplatelet therapy. DLP: dyslipidemia; DM: diabetes mellitus; HT: hypertension; ITE: individual treatment effect.




## Discussion

### Principal Findings

We estimated the causal influences of risk factors associated with stroke outcomes using multiple approaches that included SPS, IPW, DRE, SCM, and mediation analysis, in addition to DML and Dragonnet NNs. Our findings indicate strong positive causal effects associated with AF and HT on stroke development, with DM exerting a weaker effect. DLP, in contrast, had little effect. Furthermore, our analysis suggests that patients with both DM and HT not currently in receipt of antiplatelet treatments would be the most likely beneficiaries of antiplatelet therapy based on the mean ITEs.

The results from the different estimators generally demonstrated consistency, although there were slight variations in specific point estimates and confidence intervals varied slightly. The estimated causal effect derived from various methods using real-world observational data is comparable with standard cohort epidemiological studies using more traditional logistic regression approaches [28,29].

### Comparison to Prior Works

SPS is a widely used method that minimizes confounding bias by adjusting baseline covariates and confounding factors and estimating treatment effects by stratum. However, SPS is sensitive to the number of strata and features that affect both treatment and outcome (confounding factors), which can lead to bias in the causal effect estimate [30-33]. In addition, some strata may be sparsely populated, making the ATE hard to define and more prone to bias [34]. Rosenbaum and Rubin [9] originally proposed dividing the strata into 5 levels and then subsequently automatically splitting the strata until the balance in the numbers of treated and control observations was achieved [25].

IPW attempts to reduce confounding of the ATE by weighting the sample with the inverse propensity score and by balancing the distribution of the covariates between the treated and untreated groups [35], thereby avoiding the problem of data sparsity that may be present in SPS, particularly with small sample sizes. However, there is a reliance on the assumption that the propensity score model correctly captures all confounding factors, which, if incorrect, may bias the ATE. Additionally, IPW is more sensitive to the model and variable selection for estimating the propensity scores, with small differences in estimated propensity scores potentially leading to large differences in estimated causal effects [36]. Finally, IPW may imprecisely estimate treatment effects if a sample size is small, leading to a propensity score close to 0 or 1 [36,37].

DRE combines propensity score and outcome regression models [38], which can lead to improvements in the robustness of model specification by allowing one of the two treatment and outcome models to be miss-specified but still provide a consistent estimation [39]. The challenge is to validly model either the propensity score or the outcome model; it may be tempting to use modern machine learning approaches or nonparametric models in DRE, but this may lead to bias if the functions are too complex, leading to overfitting [40,41]. DML was developed to address the bias from regularization and overfitting in estimating the parameter of interest, which arises when naively inserting machine learning estimators into the estimation equation. This approach consists of two critical components: (1) the use of Neyman-orthogonal moments or scores to estimate the parameters and (2) the application of cross-fitting, which provides an efficient form of data-splitting. By using both elements, DML minimizes the impact of regularization bias and overfitting on parameter estimation; this also extends to nonparametric models [14].

Applying POMs (eg, SPS, IPW, DRE) relies heavily on the assumption that the treatment assignment is independent of the PO given the observed covariates, which is known as “unconfoundedness” or the conditional independence assumption. If this assumption does not hold, the estimated causal effect will be biased. In contrast, SCMs facilitate the modeling of complex relationships between multiple causes and effects in the presence of latent or unobserved variables [4,42]. In addition, SCMs can be considered as counterfactual predictions of interventions, which can be useful in applications such as causal inference in experimental or observational studies [43-46]. However, SCMs are limited by the assumption of independence between variables and may require conceptualized causal relationship mechanisms.

The benefit of using NNs to estimate causal effects is their flexibility and power to handle high-dimensional and complex data. Shalit et al [17] introduced TARNet by sharing information between the PO of treatment and control groups, which is different from the previous model that separated the training data. More recently, Dragonnet was developed by combining propensity scores with targeted regularization, resulting in more accurate inference [18]. Dragonnet is considered more robust with very low or high propensity scores but has several limitations including sensitivity to choice of architecture and hyperparameters, dealing with only a single set of features at a time, and difficulty of interpretation [18]. Despite some limitations, Dragonnet’s benefits surpass these drawbacks, making it an attractive approach for estimating causal effects in complex real-world data.

## Strengths and Limitations

A critical aspect of causal inference, particularly in estimating CATE, involves certain assumptions, notably ignorability and positivity. Strong ignorability necessitates the observation and adjustment for all confounding variables that influence both the treatment and the outcome, while positivity ensures that every patient has a nonzero probability of receiving each treatment. In our study, we believe these assumptions are reasonably satisfied. We included a comprehensive set of covariates, such as age, sex, BMI, chronic kidney disease, and relevant comorbidities (HT, DM, DLP, and AF), which are well-documented factors influencing stroke risk and treatment decisions. However, we acknowledge that there might be unmeasured confounders not captured in our dataset. Regarding the decision on antiplatelet drug administration, we utilized detailed patient records from Ramathibodi Hospital, ensuring a thorough assessment of factors influencing treatment. Nonetheless, we recognize the potential for residual confounding and the inherent limitations of observational data. Future studies could benefit from incorporating more granular clinical data and leveraging advanced causal discovery methods to further validate these assumptions.

Causal effects can vary between individuals, which necessitates the estimation of ITEs. Treatment effects can vary between individual patients; therefore, applying a single treatment effect as CATE to all individual patients is inappropriate [47,48] as some patients may gain more or less benefit from treatments. Thus, the estimation of ITE to identify at-risk patients most likely to benefit from treatment is a major goal for stratified and precision medicine approaches. Estimating ITEs requires larger sample sizes, as individual-level estimates are less precise than aggregate-level estimates [49]. A covariate shift may result from unobserved counterfactual data but this is minimized using a weighted split-CQR approach [23].

We believe that the clinical implications of our study are significant, as understanding the causal relationships and individual treatment effects of stroke risk factors can directly influence patient care by providing more precise and personalized risk assessments. Additionally, we can conduct reviews and quality assessments of current patients in the clinic to determine who should receive further treatment. These methods enable clinicians to identify high-risk patients who would benefit most from targeted interventions, like antiplatelet therapy, thereby optimizing treatment strategies and improving patient outcomes. The use of real-world data ensures that our findings apply to everyday clinical practice.

Our study has some limitations. First, we used real-world data rather than RCT data, thus some important covariates were not previously planned, measured, and collected as part of routine clinical evaluation and were therefore unavailable for ITE estimation. Second, we acknowledge the possibility of unmeasured confounders in the observational dataset. Future studies could benefit from incorporating more granular clinical

data, such as detailed medication records, laboratory results, and lifestyle factors, to mitigate potential confounding. Third, the models used for estimating ITEs were trained and validated in only a single setting, thereby limiting their generalizability. Future research should focus on validating the models in diverse settings with different patient populations or hospitals. This external validation would help to determine whether the models' predictive performance and the estimated ITEs hold true across various contexts.

## Conclusion

This study provides comprehensive causal estimates of AF, HT, DLP, and DM on stroke using various advanced statistical and machine learning methodologies. The consistent results across multiple analytical approaches and this study's alignment with a standard cohort study reinforce the robustness of our findings. AF and HT emerged as significant risk factors for stroke, with DM showing a moderate effect, while DLP had minimal impact. Notably, the use of Dragonnet and conformal inference techniques allowed us to accurately estimate ITEs, highlighting that several high-risk patients who did not take antiplatelets at the time of data recorded, particularly those with DM or DM combined with HT, could potentially benefit from antiplatelet therapy. This suggests that personalized treatment strategies could be pivotal in reducing stroke risk among these patients.

The findings underscore the significance of individualized risk assessment and treatment personalization in clinical settings. Future research should focus on integrating these advanced causal inference models into routine clinical practice to enhance treatment outcomes for high-risk stroke patients. Additionally, the use of real-world data provides valuable insights but also presents challenges related to unmeasured confounding and data quality. Addressing these challenges in future studies will be crucial for advancing our understanding and improving stroke management strategies.

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Authors' Contributions

This study has been conceptualized by SL and AT. SL performed data management, model construction, and analysis. The manuscript was drafted by SL and revised by GJM, JA, and AT. All authors approved the final version of this manuscript.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Nested approach for interval estimates of individual treatment effect algorithm. α=.05 to cover 95% confidence interval.
[DOCX File, 21 KB - cardio_v9i1e50627_app1.docx ]

Multimedia Appendix 2
Descriptive analysis of features between stroke and nonstroke.
[DOCX File, 23 KB - cardio_v9i1e50627_app2.docx ]

Multimedia Appendix 3
Estimated causal effect from estimators. Numbers indicate conditional average treatment (risk) effect (CATE) with 95% confidence interval. * Heart disease ** top quintile low-density lipoprotein (LDL).
[DOCX File, 22 KB - cardio_v9i1e50627_app3.docx ]

Multimedia Appendix 4
Sample of 50 individual treatment effects with 95% confidence intervals and stroke risk reduction who had not received antiplatelet treatment, demonstrating the potential benefits had they been given the medication. In this plot, 3 of the samples (6%) demonstrate that a considerable number of patients could have experienced a positive impact on their stroke risk reduction had they received the antiplatelet treatment. The y-axis displays the treatment effect, while the x-axis represents each individual patient in the sample.
[DOCX File, 59 KB - cardio_v9i1e50627_app4.docx ]

## References

1. Krishnamurthi RV, Ikeda T, Feigin VL. Global, regional and country-specific burden of ischaemic stroke, intracerebral haemorrhage and subarachnoid haemorrhage: a systematic analysis of the global burden of disease study 2017. Neuroepidemiology 2020;54(2):171-179. [doi: 10.1159/000506396] [Medline: 32079017]
2. Boehme AK, Esenwa C, Elkind MSV. Stroke risk factors, genetics, and prevention. Circ Res 2017 Feb 3;120(3):472-495. [doi: 10.1161/CIRCRESAHA.116.308398] [Medline: 28154098]

3. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011 May;46(3):399-424. [doi: 10.1080/00273171.2011.568786] [Medline: 21818162]
4. Pearl J. Causality: Cambridge University Press; 2009.
5. Rubin DB. Causal inference using potential outcomes. J Am Stat Assoc 2005 Mar;100(469):322-331. [doi: 10.1198/016214504000001880]
6. Markus KA. Causal effects and counterfactual conditionals: contrasting Rubin, Lewis and Pearl. Econ Philos 2021 Nov;37(3):441-461. [doi: 10.1017/S0266267120000437]
7. Pearl J. An introduction to causal inference. Int J Biostat 2010 Feb 26;6(2). [doi: 10.2202/1557-4679.1203] [Medline: 20305706]
8. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol 1974;66(5):688-701. [doi: 10.1037/h0037350]
9. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70(1):41-55. [doi: 10.1093/biomet/70.1.41]
10. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat 1985 Feb;39(1):33-38. [doi: 10.1080/00031305.1985.10479383]
11. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc 1984 Sep;79(387):516-524. [doi: 10.1080/01621459.1984.10478078]
12. Rosenbaum PR. Model-based direct adjustment. J Am Stat Assoc 1987 Jun;82(398):387-394. [doi: 10.1080/01621459.1987.10478441]
13. Hernán M, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 2000 Sep;11(5):561-570. [doi: 10.1097/00001648-200009000-00012] [Medline: 10955409]
14. Chernozhukov V, Chetverikov D, Demirer M, Duflo E, Hansen C, Newey W. Double/debiased/Neyman machine learning of treatment effects. Am Econ Rev 2017 May 1;107(5):261-265. [doi: 10.1257/aer.p20171038]
15. Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. J Am Stat Assoc 1994 Sep;89(427):846-866. [doi: 10.1080/01621459.1994.10476818]
16. Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. Am J Epidemiol 2011 Apr 1;173(7):761-767. [doi: 10.1093/aje/kwq439] [Medline: 21385832]
17. Shalit U, Johansson FD, Sontag D. Estimating individual treatment effect: generalization bounds and algorithms. Presented at: The 34th International Conference on Machine Learning; Aug 8-11, 2017; Sydney, Australia.
18. Shi C, Blei D, Veitch V. Adapting neural networks for the estimation of treatment effects. Presented at: Advances in Neural Information Processing Systems 32 (NeurIPS 2019); Dec 8-14, 2019; Vancouver, Canada.
19. Curth A, Schaar M. Nonparametric estimation of heterogeneous treatment effects: from theory to learning algorithms. Presented at: The 24th International Conference on Artificial Intelligence and Statistics; Apr 13-15, 2021.
20. Lei L, Candès EJ. Conformal inference of counterfactuals and individual treatment effects. arXiv. Preprint posted online on Jun 11, 2020. [doi: 10.48550/arXiv.2006.06138]
21. Vovk V, Gammerman A, Shafer G. Algorithmic Learning in a Random World: Springer; 2005.
22. Shafer G, Vovk V. A tutorial on conformal prediction. J Mach Learn Res 2008;9(3):371-421. [doi: 10.1145/1390681.1390693]
23. Tibshirani RJ, Foygel Barber R, Candes E, Ramdas A. Conformal prediction under covariate shift. Presented at: Advances in Neural Information Processing Systems 32 (NeurIPS 2019); Dec 8-14, 2019; Vancouver, Canada.
24. Sharma A, Kiciman E. DoWhy: an end-to-end library for causal inference. arXiv. Preprint posted online on Nov 9, 2020. [doi: 10.48550/arXiv.2011.04216]
25. Imbens GW, Rubin DB. Causal Inference in Statistics, Social, and Biomedical Sciences: Cambridge University Press; 2015. [doi: 10.1017/CBO9781139025751]
26. Blöbaum P, Götz P, Budhathoki K, Mastakouri AA, Janzing D. DoWhy-GCM: an extension of dowhy for causal inference in graphical causal models. arXiv. Preprint posted online on Jun 14, 2022. [doi: 10.48550/arXiv.2206.06821]
27. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol 2013 Oct;42(5):1511-1519. [doi: 10.1093/ije/dyt127] [Medline: 24019424]
28. O'Donnell MJ, Chin SL, Rangarajan S, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. Lancet 2016 Aug;388(10046):761-775. [doi: 10.1016/S0140-6736(16)30506-2]
29. Tirschwell DL, Smith NL, Heckbert SR, Lemaitre RN, Longstreth WT Jr, Psaty BM. Association of cholesterol with stroke risk varies in stroke subtypes and patient subgroups. Neurology 2004 Nov 23;63(10):1868-1875. [doi: 10.1212/01.wnl.0000144282.42222.da] [Medline: 15557504]
30. Imai K, Dyk DA. Causal inference with general treatment regimes: generalizing the propensity score. J Am Stat Assoc 2004;99(467):854-866. [doi: 10.1198/016214504000001187]
31. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods 2010 Sep;15(3):234-249. [doi: 10.1037/a0019623] [Medline: 20822250]

32. Hullsiek KH, Louis TA. Propensity score modeling strategies for the causal analysis of observational data. Biostatistics 2002 Jun;3(2):179-193. [doi: 10.1093/biostatistics/3.2.179] [Medline: 12933612]
33. Adelson JL, McCoach DB, Rogers HJ, Adelson JA, Sauer TM. Developing and applying the propensity score to make causal inferences: variable selection and stratification. Front Psychol 2017;8:1413. [doi: 10.3389/fpsyg.2017.01413] [Medline: 28861028]
34. Morgan SL, Winship C. Counterfactuals and Causal Inference: Methods and Principles for Social Research: Cambridge University Press; 2014.
35. Brookhart MA, Wyss R, Layton JB, Stürmer T. Propensity score methods for confounding control in nonexperimental research. Circ Cardiovasc Qual Outcomes 2013 Sep 1;6(5):604-611. [doi: 10.1161/CIRCOUTCOMES.113.000359] [Medline: 24021692]
36. Chesnaye NC, Stel VS, Tripepi G, et al. An introduction to inverse probability of treatment weighting in observational research. Clin Kidney J 2021 Aug;15(1):14-20. [doi: 10.1093/ckj/sfab158] [Medline: 35035932]
37. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol 2008 Sep 15;168(6):656-664. [doi: 10.1093/aje/kwn164] [Medline: 18682488]
38. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. Biometrics 2005 Dec;61(4):962-973. [doi: 10.1111/j.1541-0420.2005.00377.x] [Medline: 16401269]
39. Robins J, Rotnitzky A. Comment on the Bickel and Kwon article, 'Inference for semiparametric models: Some questions and an answer'. Stat Sin 2001;11:920-936.
40. Benkeser D, Carone M, Laan MJVD, Gilbert PB. Doubly robust nonparametric inference on the average treatment effect. Biometrika 2017 Dec;104(4):863-880. [doi: 10.1093/biomet/asx053] [Medline: 29430041]
41. Zivich PN, Breskin A. Machine learning for causal inference: on the use of cross-fit estimators. Epidemiology 2021 May 1;32(3):393-401. [doi: 10.1097/EDE.0000000000001332] [Medline: 33591058]
42. Pearl J. Causal diagrams for empirical research. Biometrika 1995 Dec;82(4):669-688. [doi: 10.2307/2337329]
43. Pearl J. Causal and counterfactual inference. In: The Handbook of Rationality: MIT Press; 2021:427-438. [doi: 10.7551/mitpress/11252.003.0044]
44. Pawlowski N, Castro D, Glocker B. Deep structural causal models for tractable counterfactual inference. Presented at: Advances in Neural Information Processing Systems 33 (NeurIPS 2020); Dec 6-12, 2020.
45. Xia K, Pan Y, Bareinboim E. Neural causal models for counterfactual identification and estimation. arXiv. Preprint posted online on Sep 30, 2022. [doi: 10.48550/arXiv.2210.00035]
46. Zhang J, Tian J, Bareinboim E. Partial counterfactual identification from observational and experimental data. Presented at: The 39th International Conference on Machine Learning; Jul 17-23, 2022; Baltimore, MD.
47. Nguyen TL, Collins GS, Landais P, Le Manach Y. Counterfactual clinical prediction models could help to infer individualized treatment effects in randomized controlled trials-an illustration with the International Stroke Trial. J Clin Epidemiol 2020 Sep;125:47-56. [doi: 10.1016/j.jclinepi.2020.05.022] [Medline: 32464321]
48. Sperrin M, Martin GP, Pate A, Van Staa T, Peek N, Buchan I. Using marginal structural models to adjust for treatment drop-in when developing clinical prediction models. Stat Med 2018 Dec 10;37(28):4142-4154. [doi: 10.1002/sim.7913] [Medline: 30073700]
49. Hoogland J, IntHout J, Belias M, et al. A tutorial on individualized treatment effect prediction from randomized trials with a binary endpoint. Stat Med 2021 Nov 20;40(26):5961-5981. [doi: 10.1002/sim.9154] [Medline: 34402094]


## Abbreviations

**AF:** atrial fibrillation
**ATE:** average treatment effect
**CATE:** conditional average treatment effect
**CKD:** chronic kidney disease
**CQR:** conformal quantile regression
**DLP:** dyslipidemia
**DM:** diabetes mellitus
**DRE:** doubly robust estimation
**HT:** hypertension
***ICD-10*:** *International Classification of Diseases, 10th Revision*
**IPW:** inverse probability weighting
**ITE:** individualized treatment effect
**NDE:** natural direct effect
**NIE:** natural indirect effect
**NN:** neural network
**PO:** potential outcome
**POM:** potential outcome model

**RCT:** randomized controlled trial
**SCM:** structural causal model
**SPS:** stratified propensity score
**XGBoost:** Extreme Gradient Boosting

*Edited by A Coristine; submitted 07.07.23; peer-reviewed by J Rivers, M Wright, N Kakaletsis, S Jaroszewicz; revised version received 23.11.24; accepted 24.11.24; published 08.01.25.*

*Please cite as:*
*Lolak S, Attia J, McKay GJ, Thakkinstian A*
*Application of Dragonnet and Conformal Inference for Estimating Individualized Treatment Effects for Personalized Stroke Prevention: Retrospective Cohort Study*
*JMIR Cardio 2025;9:e50627*
*URL: https://cardio.jmir.org/2025/1/e50627*
*doi:10.2196/50627*

# Electronic Clinical Decision Support System for Stroke Risk Screening in Patients With Atrial Fibrillation in Mental Health Care: Mixed Methods Study

Dina Farran[1], MPH, PhD; Hou Wang Cheang[1], MSc; Juliana Onwumere[2,3,4], PhD; Mark Ashworth[5†], DM; Fiona Gaughran[1,2], MD

[1]Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, 16 De Crespigny Park, London, United Kingdom

[2]South London and Maudsley NHS Foundation Trust, London, United Kingdom

[3]NIHR Biomedical Research Centre for Mental Health South London and Maudsley NHS, London, United Kingdom

[4]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[5]School of Life Course and Population Sciences, King's College London, London, United Kingdom

[†]deceased

**Corresponding Author:**
Dina Farran, MPH, PhD
Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, 16 De Crespigny Park, London, United Kingdom

## Abstract

**Background:** Electronic clinical decision support systems (eCDSSs) are key to the digital transformation of health care. Despite their growing adoption, little is known about the perspectives of mental health clinicians on the implementation of eCDSS to assist them in managing physical health conditions within mental health care settings.

**Objective:** This study aimed to explore how clinicians in older adult mental health services manage stroke risk in patients with atrial fibrillation (AF) and comorbid serious mental illness who are admitted to the hospital under their care. It also sought to examine clinicians' views on the potential role of an eCDSS in enhancing stroke risk assessment and management.

**Methods:** A cross-sectional mixed methods study was conducted between March and May 2023 in 3 inpatient wards for mental health of older adults at South London and Maudsley NHS (National Health Service) Foundation Trust. Health care professionals, including psychiatrists and pharmacists, participated in a web-based survey and individual semistructured interviews. Ethical approval and informed consent were obtained. A descriptive analysis was conducted on the survey data, while interview data were analyzed thematically using an inductive approach.

**Results:** In total, 10 clinicians participated in the study. Thematic analysis revealed 2 primary themes. First, clinicians reported significant challenges in clinical practice, including difficulties accessing patient medical histories, limited expertise in managing physical health conditions, fragmented care pathways, and the impact of mental health symptoms such as psychotic beliefs on stroke prevention. Second, clinicians identified strategies to improve practice, such as embedding alerts in electronic records, establishing clear organizational policies, and providing tailored training on AF-related stroke management. Clinicians recognized the potential of an eCDSS to enhance clinical effectiveness, improve the identification of high-risk patients, ensure safer and more consistent care, and save time. However, they expressed concerns about potential risks, including rigidity in decision-making, overreliance on the tool, false positives, reduced critical thinking, annoyance, and increased workload.

**Conclusions:** This study highlights the challenges and opportunities in managing AF-related stroke risk in mental health settings. While clinicians acknowledged the potential of an eCDSS to improve care quality and efficiency, addressing concerns about its design and implementation is essential. These insights can inform the development of eCDSS tools that effectively balance benefits with user needs, ultimately improving patient outcomes in mental health services.

***(JMIR Cardio 2025;9:e66428)*** doi:10.2196/66428

**KEYWORDS**

atrial fibrillation; mental illness; stroke risk; stroke; clinical decision support systems; CDSS; digital health alerts; decision support; heart; cardiac; arrhythmia; cardiology; interview; qualitative approach; thematic analysis; experience; attitude; opinion; perception; perspective; medical informatics

## Introduction

Electronic clinical decision support systems (eCDSSs) are software-based tools that analyze patient data locked in electronic health records (EHRs) and provide clinicians with relevant clinical support in the form of alerts or reminders [1]. Given the increasing volume of clinical information and the rapid advances in the field of medicine, eCDSSs can be pivotal in providing evidence-based clinical guidelines and tailored clinical support with personalized guidance for diagnostic, therapeutic, and preventive interventions [1].

eCDSSs have gained substantial attention in recent years for their potential to assist health care professionals in selecting appropriate treatment, managing medication (eg, dosing, contraindications, potential interactions, and side effects), calculating risk scores, identifying patients at risk, tracking patient progress over time, and documenting clinical data [2]. This has the potential to reduce medical errors and enhance health outcomes. However, eCDSSs can also have drawbacks. Alert fatigue can result in health care professionals becoming desensitized to notifications and potentially missing important information. This is often the case when the digital tool is overused or poorly designed [2-4]. Clinicians can report feeling overwhelmed with the volume and frequency of alerts, which may, in turn, disrupt workflow, resulting in less face-to-face time with patients [2-4]. Additionally, eCDSSs can lead to incorrect recommendations if the data input are inaccurate or of poor quality [2-4].

Many eCDSSs have been developed to help health care professionals manage physical health conditions, including atrial fibrillation (AF) and associated stroke risk [5]. AF is an arrhythmia characterized by irregular heartbeats. AF disrupts the ability of the heart to pump blood effectively, resulting in a higher risk of blood clotting within the left atrium of the heart and an increased risk of stroke [6]. Based on the National Institute for Health and Care Excellence (NICE) guidelines, patients with AF should undergo a stroke and bleeding risk assessment using the $CHA_2DS_2$-VASc and ORBIT (Outcomes Registry for Better Informed Treatment of Atrial Fibrillation) scales, respectively. NICE recommends oral anticoagulation (OAC) therapy for patients with a $CHA_2DS_2$-VASc score ≥2 and asks clinicians to consider anticoagulation for males with a $CHA_2DS_2$-VASc of 1. When the bleeding risk is low (ORBIT score <3), OAC therapy can be initiated or continued; however, when the risk is moderate or high, careful consideration of the benefits and potential risks associated with the therapy is required [7].

Research assessing the prevalence of AF among people with mental disorders is scarce. A recent nationwide population-based study reported that the risk of AF increased by 2-fold in patients with bipolar disorder or schizophrenia and by 1.5 - to 1.7-fold in those with depression, insomnia, and anxiety disorders compared to controls [8]. Additionally, people living with a mental illness are at increased risk of cardiovascular disease (including strokes), mainly due to risk factors such as obesity, smoking, diabetes, hypertension, and dyslipidemia [8]. Despite evidence supporting the benefits of OAC therapy, people with AF and comorbid mental health conditions are less likely than the general population to be prescribed OAC therapy [9].

While many studies have evaluated the feasibility, acceptability, and effectiveness of eCDSSs in supporting the management of AF and related stroke risk in general acute hospital settings, these studies were not conducted in mental health care settings [10-19]. Implementing an eCDSS that screens for the stroke risk among patients with AF admitted to a mental health hospital is key to early prevention and quality of life improvement.

This study focuses on older adult mental health services to investigate how AF-related stroke risk is managed in individuals with comorbid severe mental illness, addressing gaps in care to improve outcomes in this high-risk population. Specifically, the study aims to explore (1) clinicians' experiences in managing AF-related stroke prevention in secondary mental health care services and (2) their perspectives on the potential impact of an eCDSS in enhancing the quality of care in these settings.

## Methods

### Design

This cross-sectional study used a mixed methods research design, incorporating a short web-based survey and individual semistructured interviews.

### Ethical Considerations

Ethical approval was granted by the King's College London Research Ethics Committee, SLaM Capacity and Capability (Trust R&D Reference: R&D2023/004) and NHS (National Health Service) Health Research Authority (22/HRA/5452). The study was conducted in accordance with the principles of the Declaration of Helsinki (1996) and all applicable regulatory requirements, including but not limited to the UK policy framework for health and social care research, Trust and Research Office policies and procedures, and any subsequent amendments. Information gathered in this study was kept confidential and managed based on the Data Protection Act, NHS Caldicott Guardian, The Research Governance Framework for Health, and Social Care and HRA Approval. Informed consent was obtained from all participants before data collection. Potential participants were provided with an information leaflet outlining the study's purpose, procedures, and their right to withdraw at any time without consequence. Participants were informed that their participation was voluntary. To ensure privacy and confidentiality, all participant data were anonymized and de-identified. Personal identifiers were removed from the transcripts, and the data were stored securely in compliance with data protection regulations. Any information that could potentially identify participants was excluded from the analysis and reporting. No compensation was provided to participants in this study, as the nature of the study did not involve any direct financial incentives for participation.

### Recruitment

The study was conducted between March and May 2023 in 3 mental health of older adult (MHOA) inpatient wards at South London and NHS Foundation Trust (SLaM). These wards provide specialized care for older adults with a range of mental

health conditions, often coexisting with physical health challenges. Purposive sampling was used to identify and recruit participants who were likely to provide clinical care for patients with AF and a comorbid mental health condition.

Senior management on potential wards were first approached by the research team and given brief information regarding the nature of the study and the eCDSS to be implemented. Wards that expressed an interest in the study were provided with further detailed information.

A subgroup of health care professionals, including psychiatrists and pharmacists working on recruited wards, were all invited to complete a short survey and take part in an individual interview. Potential participants varied in terms of professional seniority and clinical experience, with a focus on including diverse perspectives to enrich the findings. Potential participants were given an information leaflet and an opportunity to ask and discuss any further concerns regarding the study. If in agreement to enroll, participants were asked to sign a consent form. The number of participants required for this study was not pre-estimated and was fully dependent on theme saturation in the qualitative part.

### Intervention

The eCDSS consists of a visual prompt integrated into the EHR, which is triggered whenever a patient with documented AF, either chronic or newly diagnosed upon admission, is admitted to the hospital. Using natural language processing, the system identifies references to AF in clinical notes and alerts clinicians to confirm the presence of AF, complete clinical assessments of stroke and bleeding risks using the $CHA_2DS_2$-VASc and ORBIT scales, and record the scores in the EHR. For patients found to have a high risk of stroke, clinicians are prompted to refer them to OAC clinics for specialized care.

### Data Collection

All participants were asked to complete a short web-based survey designed to gather demographic and professional background information, including their age, gender, professional background, and years of clinical experience. The questionnaire was developed to assess clinicians' awareness and confidence regarding AF-related stroke prevention. It included a series of statements related to their knowledge of AF guidelines, their confidence in assessing stroke and bleeding risks using the $CHA_2DS_2$-VASc and ORBIT scales, and their confidence in managing patients at risk of stroke. Each statement was rated on a Likert scale, ranging from strongly disagree to strongly agree (Multimedia Appendix 1). Example statements included the following: "I am confident in identifying atrial fibrillation patients eligible for oral anticoagulation therapy," "I am confident in assessing the stroke risk using the $CHA_2DS_2$-VASc tool and the bleeding risk using the ORBIT tool," and "I am confident in managing atrial fibrillation-related stroke risk in mental healthcare settings."

The development of the questionnaire involved collaboration with field experts, including psychiatrists, general practitioners (GPs), and health care professionals with expertise in stroke prevention and mental health care. This ensured that the items were relevant to the clinical context and aligned with current guidelines for stroke prevention in patients with AF. The questionnaire was pretested with a small sample of health care professionals to ensure clarity and relevance, and minor revisions were made based on their feedback.

In addition to the survey, an interview schedule was created to explore participants' experiences with AF-related stroke prevention in secondary mental health care services and the potential impact of an eCDSS on clinician-led care in MHOA wards. The interview topic guide was informed by feedback from field experts, ensuring that the questions were comprehensive and aligned with the research objectives (Multimedia Appendix 2).

Participants were contacted via email, and interviews were scheduled according to their availability. The interviews were conducted via Microsoft Teams, with the same researcher (DF) leading all interviews. Each interview lasted approximately 20 minutes and followed a semistructured format with key prompts to direct the discussion while allowing flexibility for participants to share their insights. All interviews were audio-recorded, transcribed verbatim, and deidentified prior to analysis.

### Data Analysis

Data collected through the questionnaire were analyzed descriptively. Responses to Likert-scale items were summarized to capture the distribution of confidence levels, perceptions of current care quality, and attitudes toward the eCDSS. Demographic and professional background data were also summarized to contextualize participants' responses.

Thematic analysis was conducted following Braun and Clarke's framework, incorporating updated guidance from their 2023 work on good practices in thematic analysis [20]. An inductive, data-driven approach was used to allow the themes to emerge directly from the data. The analysis was conducted by 2 members of the research team (DF and HC) and involved several iterative steps. First, both researchers immersed themselves in the data by reading and rereading the transcripts to gain a comprehensive understanding of the content. Descriptive codes were independently generated for each transcript by both researchers, with codes refined and adjusted during subsequent readings. A coding framework was then collaboratively developed based on the descriptive codes, and this framework was iteratively revised to ensure alignment and accommodate different perspectives.

Codes were grouped into broader themes that reflected significant patterns in the data and addressed the research questions. This process involved exploring similarities and differences within and across transcripts and examining patterns based on participant characteristics. Themes were iteratively refined, defined, and labeled to ensure clarity, coherence, and alignment with the data. To enhance the credibility of the findings, the themes were discussed with clinical experts and refined based on their input.

### Reflexivity and Methodological Rigor

The research team actively engaged in reflexivity throughout the study to address potential biases. One of the researchers

(DF) brought substantial expertise in applied health informatics and the clinical implications of stroke prevention, which could have shaped their perspective during data collection and analysis. To mitigate this, the researcher frequently reflected on preconceptions and assumptions, documented potential biases, and maintained an ongoing record of decisions made during the research process. Additionally, a second researcher (HC) independently conducted parallel analyses to provide an alternative lens and challenge interpretations.

Transparency and rigor were further enhanced by explicitly acknowledging positionality and engaging clinical experts in the refinement of themes. Methodological rigor was addressed by ensuring confirmability through the maintenance of a clear audit trail of the research process. Dependability was established through the use of a systematic and replicable analytic approach. Credibility was supported by triangulation between researchers and consultation with clinical experts, while transferability was facilitated by providing a detailed description of the study context and participants to enable readers to assess the applicability of the findings to similar settings.

## Results

The sample comprised 10 participants (from a total of 15 invited clinicians), of whom 6 reported their gender as female and 4 as male. Participants' ages ranged between 25 and 46 years, with a mean age of 32 years. In terms of professional background, a slightly larger number were psychiatrists, which included 3 participants at the consultant level and 3 at a more junior level. The remaining participants (n=4) were pharmacists. The mean years of clinical experience (defined as years a health care professional has spent in clinical practice since professional qualification) was 7.25.

In total, 50% of participants (n=5) considered that AF-related stroke prevention is suboptimal on the wards where they work. Half of participants reported being confident or somewhat confident in managing AF-related stroke prevention in mental health care settings or in making referrals to OAC clinics. Around 60% reported being confident or somewhat confident using the $CHA_2DS_2$-VASc tool to assess the risk of stroke, whereas only 30% reported being confident or somewhat confident using the ORBIT tool to assess the risk of bleeding. Almost all participants strongly agreed that having access to an eCDSS would help them to better assess stroke and bleeding risks in patients with AF.

Thematic analysis of the interviews identified 2 overarching themes related to prevention of AF-related stroke: (1) challenges faced on wards and (2) strategies needed to improve practice (Figure 1). As for the potential impact of an eCDSS in improving quality of care, two themes emerged: (1) perceived benefits and (2) perceived risks (Figure 2).

**Figure 1.** Coding tree of themes pertaining to the prevention of AF-related stroke in mental health of older adult (MHOA) ward.

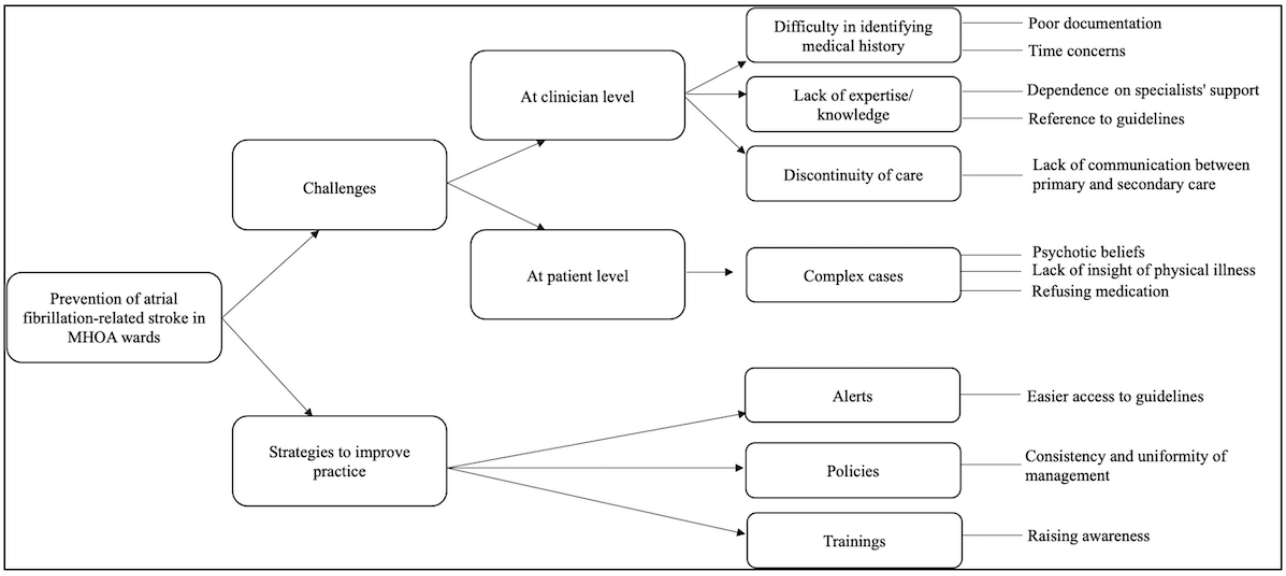
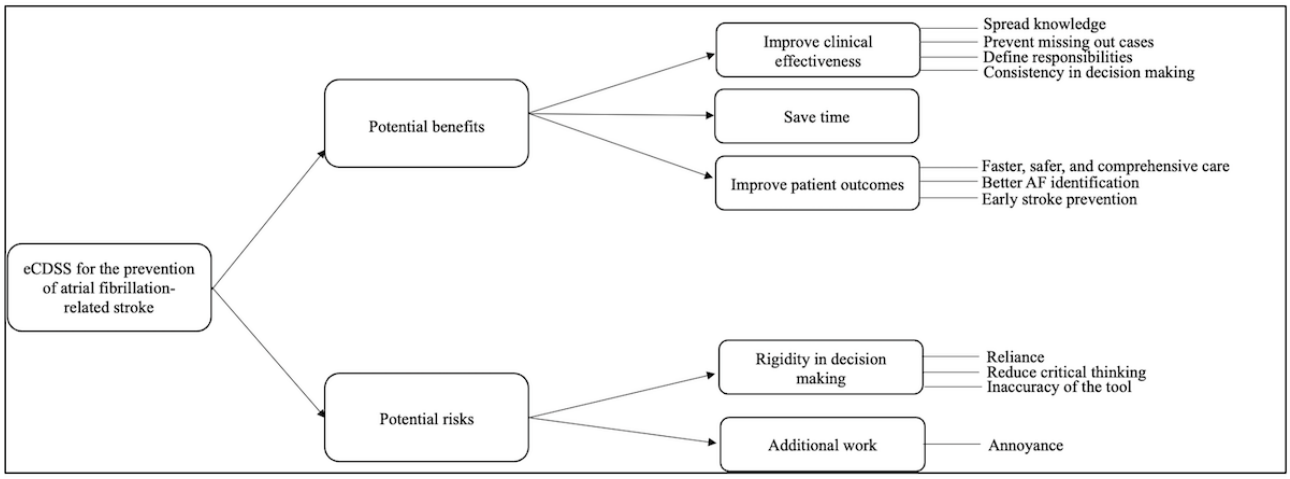


**Figure 2.** Coding tree of themes pertaining to the potential impact of an electronic clinical decision support system (eCDSS) for the prevention of atrial fibrillation (AF)–related stroke in mental health of older adult wards.




## Prevention of AF-Related Stroke in MHOA Wards

### *Challenges*

Participants discussed challenges in the prevention of AF-related stroke in MHOA wards at 2 levels.

At the clinician level, many participants reported that identifying a medical history of AF from the electronic clinical notes is a challenging and time-consuming task. Some of them attributed this to poor documentation of physical health conditions in mental health care settings. Another challenge is clinician lack of knowledge and expertise in the management of physical conditions. To optimize management of physical long-term conditions, most clinicians would seek support from specialists or refer to guidelines such as NICE. Discontinuity of care provision and lack of communication between primary and secondary care were also considered obstacles in the prevention of AF-related stroke in MHOA wards. Participants expressed their concerns about the lack of coordination and follow-up with GPs and its effect on the quality of care (Figure 1 and Table 1).

**Table .** Illustrative quotations for the identified subthemes.

| Subthemes | Examples |
|---|---|
| Difficulty in identifying medical history | • "sometimes it's a bit more difficult to get the appropriate history like have they had any previous strokes, do they have any comorbidities, or do they have any other risk factors, family history of stroke."<br>• "The recording of ECG in the department sometimes isn't great. Uploading ECG onto our clinical document system doesn't happen a lot of the times, so one of the challenges is to find the ECG initially for people to identify it."<br>• "the doctor usually will read the clinical notes to find out the cardiac history for the patient. This is how we usually identify AF or any stroke history for patient. I guess the challenge is that it's a bit time wasting because the doctors will have to review previous clinical letters or any discharge summary." |
| Lack of expertise or knowledge | • "I don't feel very confident at all, to be honest. I did my foundation years ago and like mental health placement before starting psychiatry training. You know, I used to deal with it a bit on the medical take, but I think if I were to identify a new AF on admission, I'd just discuss it with specialists. But independently, if you were like, sort this out yourself I wouldn't feel confident."<br>• "Sadly not knowledgeable, I would immediately go and look up the NICE guidance to see the most up to date guidelines because we don't use it all the time. I probably know when to be worried. I would know where to find the information. But it wouldn't be all in my mind. But I wouldn't say I'm knowledgeable at all."<br>• "Well, I'm aware of the CHA2DS2VASc and ORBIT, but not familiar with their use."<br>• "I'd speak to my medical colleagues or end up looking at guidelines and trying to follow because I wouldn't be kind of regularly checking on what the latest guidance is. So it'd be something I'd have to refresh myself when the situation comes up." |
| Discontinuity of care | • "I also think the communication between primary and secondary care is probably one of the biggest obstacles."<br>• "The main obstacle I think is the discontinuity of care between secondary care and primary care and also from our side we're a bit of a like mental health setting so when we're starting medications that might increase the bleed risk, I think that's something that we don't have much of a process for here." |
| Complex cases | • "Patients refuse to take medications because of their psychotic beliefs or them just having given up, depression, wanting to die, basically, so they don't have the will to get better."<br>• "A lot of people at this stage of their dementia lack insight to their physical health."<br>• "If we assess the patient to have a high risk of stroke and we want to start on anticoagulants, a lot of our patients actually refuse medication."<br>• "Patients don't want any of the medication either. They think we're trying to poison them."<br>• "Patients not taking their medication is quite a common scenario on my ward." |

| Subthemes | Examples |
|---|---|
| Alerts | • "having PJS alerts is very key, so something that would prompt people and you know provide a very easy pathway for them to follow the guidance rather than kind of spending time to look things up and then not knowing if it's accurate or if it's appropriate."<br>• "Reminding people and having these tools easily accessible, you know, so they don't have to look them up so that they're incorporated probably in the notes."<br>• "So I think like very concise and clear guidelines. And probably like a hyperlink to where you can do the CHA2DS2VASc and ORBIT scoring and it would then maybe have the kind of action points for the outcome scores.<br>• "I think more information. I think what would be good obviously is an explanation on why that's being recommended, just so that physicians are aware. It'd be great if they could say like well, if the CHA2DS2VASc is greater than this much in these patients, we recommend that, and for example, if it is recommended that they just get monitored annually. Just a statement saying why."<br>• " So again I I guess the hardest thing with the AF question is that you end up with the CHA2DS2VASc and ORBIT scores or whatever, which is fine. However, you might end up with somebody who you know they score high on CHA2DS2VASc, but then high for ORBIT as well. They obviously have something that flags on, but you don't really know what that means. There's still a judgment involved, I guess. So if it was to help with making that part of the Judgment, that I guess would be helpful."<br>• "I think it would be worth including the e-mail address or way to refer to cardiology and everything you need to action that request. What I think would be most useful is kind of like how to do it, what you're supposed to do with it and how to do what you're supposed to do with it?"<br>• "Well, obviously, where to refer if you need to. Where to refer if you need support or help. And maybe also some contact for local OAC clinics and know where to refer people depending on their GP or their home address or their hospital location. Just so we're not kind of running around and we just got kind of single referral point of access." |
| Policies | • "maybe just having like policies on how to manage AF and sort of guidance. I know we have like the physical health guidelines here, but yeah, like a clear pathway would be great."<br>• "I think having some uniformity of how we address things across different wards would be helpful and would provide consistency."<br>• "at trust level, I guess having a policy. Because at SLaM we have a bit of a problem in that we don't have policies for general health conditions, so I always get calls about where's SLaM's policy for that specific physical health conditions." |
| Trainings | • "At individual level, probably training for our junior doctors and the rest of the team as well. Yeah, training on kind of recognition, and latest guidelines. Also as I mentioned, kind of what to do in a typical situation which you know we might come into contact with our patient cohort."<br>• "I guess it would be just kind of ongoing training to make sure that we are up to date with guideline changes and things."<br>• "I think improving clinicians knowledge of how to manage it. So more awareness of when anticoagulant has to be indicated, how to manage people on anticoagulants. So yeah, just knowledge and like teaching sessions would be great."<br>• "Perhaps during the induction process, this is one of those things that Drs have to be inducted in the expectation that these are the steps that we need to take if somebody does have AF." |
| Improve clinical effectiveness | |

| Subthemes | Examples |
| --- | --- |
| | • "And then another benefit would be increasing awareness in clinicians about strokes and prevention of strokes especially in elderly patients where these are more common."<br>• "Prompting clinicians and also alerting them can make people feel comfortable, knowing that they've got, like, some sort of system in place and like everyone know where the responsibilities lie in terms of managing."<br>• "I mean, I think if it's consistent, if it's done for all patients, then we'll pick up more patients or less will be missed whether or not patient gonna be compliant is the different thing. But at least we'll pick them up and an attempt to sort of preventing stroke will be made." |
| Save time | • "The benefits would be to identify, you know information that we want to quickly."<br>• "I guess because it's hard for clinicians to keep track of the patients or like all the patients all the time, I guess it will help to speed up like to help their job a little bit. When they get notified, they can further look into it rather than missing it out completely."<br>• "the benefits can help you achieve something or kind of assessment risks and benefits and things a bit more quickly."<br>• "Then, certainly it would sort of take into account all the guidelines at the same time and point you in the right direction, which just makes you save time and effort." |
| Improve patient care | • "Starting prevention and treatment earlier."<br>• "The benefits would be that people are appropriately anticoagulated and we avoid strokes especially that we've got lots of people with kind of high physical health care morbidities and vascular risk factors."<br>• "Obviously I think it will reduce the number of patients who might not be getting the appropriate treatment or the appropriate prevention, so that would be the main benefit."<br>• "I guess they can help to prioritise workload for them and it will also highlight physical health problems and I think it will help them to make the decision with a more like a well-rounded approach like considering the physical health factors, not just the mental health."<br>• "So enabling better patient care, faster, maybe more comprehensive, maybe just safer basically if it's flagging things up."<br>• It "will improve the safety in a tremendous amount to be honest" |
| Rigidity in decision-making | |

| Subthemes | Examples |
|---|---|
| | • "I think the main thing is that people can just become kind of blinkered or rigid in their decision-making and kind of forget about the the specific individual factors for that patient that may be quite relevant, but don't necessarily come up on the on the tool."<br>• "we can start to think that's the only thing that matters. So like with AF preventing stroke they might just care about the CHA2DS2VASc and ORBIT scores and see what the decision tool makes and they might not be looking at what other things are happening with the patients"<br>• "as long as it's a suggestion and It's not going to prevent sort of clinical decision-making, it is fine. And I think we need to make sure that yes, it is a prompt and everything but at the end of the day the clinician has to make a decision based on what they believe is appropriate even if it's not exactly what the tool says. It should be fine as long as we don't take the thinking out of it and it's sort of like a tool rather than mandatory in a sense."<br>• "You know automated system can never replace a human you know, because the human person takes into account the individual with their specific circumstances. So most people will probably fit into that system, but there will be others that require more individualized approach."<br>• "there may be mistakes I guess from the electronic system and identifying the wrong thing or misleading us. And I worry that maybe at some point clinicians will just think that if it's not been highlighted to me electronically, I don't need to think about it. I think there's always a danger of that for anything so."<br>• "if the electronic system has any fault to it, then they could potentially lead to a mess."<br>• "Uh harms of this system would be over reliance of electronic systems, we can become a bit over relying I think. A bit of an overreliance sometimes isn't great."<br>• "It might cause dependency. Clinicians could be just relying on the screening of the electronic system rather than themselves reading into the history."<br>• "I think maybe the disadvantages are that clinicians will be relying on electronic decision support system rather than thinking for themselves or trying to find their information" |
| Additional work | • "There's lots of things already that we have to do on ePJS and another form is likely, unless it's really prompting, it's likely to get forgotten and avoided actively or found to be quite annoying."<br>• "The harm is I don't know how the tool is. If it's going to pick up, if it's going to be accurate and picking up what it picks up, if it's going to end up more work for the NHS because they're scrolling through lots of data."<br>• "I guess maybe more paperwork."<br>• "because there's no more time in the day, you know, like there are sort of limits within which these things are being introduced and It's like when you're filling in that new form you are not doing something else and whatever that may be." |

At the patient level, patients with mental illness admitted to MHOA wards are complex, generally having both physical and mental health diagnoses. Illness-related symptoms (eg, delusional beliefs) and active features of illness may result in patient denial of being physically ill, saying that they want to die or refusing medication (Figure 1 and Table 1).

### *Strategies to Improve Practice*

To improve AF-related stroke prevention in MHOA wards, most participants suggested sending alerts to clinicians on the patient EHR containing the latest guidelines, including tools for stroke and bleeding risk assessment, guidance on how to interpret the scores, and guidance on how to refer patients at high risk of stroke to OAC clinics. Although most of the information is available online, health care professionals highlighted the importance of making it easily accessible when needed to increase efficiency. They also suggested having policies at the system level for AF-related stroke management to ensure consistency and uniformity in health care provision. At a more individual level, training sessions for health care professionals on the management of AF and how to perform stroke and bleeding risk assessments based on the latest guidelines were thought helpful (Figure 1 and Table 1).

- Subject 9 (psychiatrist): "So I think like very concise and clear guidelines. And probably like a hyperlink to where

you can do the $CHA_2DS_2$-VASc and ORBIT scoring and then maybe have the kind of action points for the outcome scores.
- Subject 4 (pharmacist): "maybe just having like policies on how to manage AF and sort of guidance. I know we have like the physical health guidelines here, but yeah, like a clear pathway would be great."

### eCDSS for the Prevention of AF-Related Stroke

#### *Potential Benefits*

Most health care professionals reported that an eCDSS for the prevention of AF-related stroke in MHOA wards would improve clinical effectiveness. This could be through spreading knowledge on the management of the condition among clinicians specialized in mental health, defining responsibilities, and ensuring consistency in decision-making. Additionally, participants emphasized the effectiveness of the tool in saving time and speeding up the clinical assessment process. They also reported that an eCDSS would be helpful in improving patient health outcomes as it will ensure faster, safer, and more comprehensive care; improve AF identification in MHOA wards; reduce the chances of getting inappropriate treatments; and ensure early stroke prevention (Figure 2 and Table 1).

- Subject 4 (pharmacist): "Prompting clinicians and also alerting them can make people feel comfortable, knowing that they've got, like, some sort of system in place and like everyone know where the responsibilities lie in terms of managing."
- Subject 3 (psychiatrist): "the benefits can help you achieve something or kind of assessment risks and benefits and things a bit more quickly."
- Subject 7 (pharmacist): "So enabling better patient care, faster, maybe more comprehensive, maybe just safer basically if it's flagging things up."

#### *Potential Risks*

While an eCDSS can be a very helpful tool for health care professionals, it may have potential risks; one of these is the rigidity in decision-making. Participants reported that they may become overreliant on such digital tools, which may influence their critical thinking skills. They also emphasized that errors in the accuracy of the tool may be misleading and could result in wrong recommendations. Participants were kind of worried about the increased workload caused by the digital tool and reported that annoyance and alert fatigue could be other downsides (Figure 2 and Table 1).

- Subject 3 (psychiatrist): "I think the main thing is that people can just become kind of blinkered or rigid in their decision-making and kind of forget about the specific individual factors for that patient that may be quite relevant, but don't necessarily come up on the on the tool."
- Subject 4 (pharmacist): "Uh harms of this system would be over reliance on electronic systems, we can become a bit over relying I think. A bit of an overreliance sometimes isn't great."
- Subject 6 (pharmacist): "if the electronic system has any fault to it, then they could potentially lead to a mess."
- Subject 2 (psychiatrist): "There's lots of things already that we have to do on ePJS and another form is likely, unless it's really prompting, it's likely to get forgotten and avoided actively or found to be quite annoying."

## Discussion

### Principal Findings

This was an exploratory study that sought to understand mental health care professionals' experience in the prevention of AF-related stroke and their perspective on the potential impact of an eCDSS in improving that experience. Clinicians reported many challenges related to stroke prevention in MHOA wards, including difficulty identifying patient pertinent medical history, perceived lack of knowledge and expertise in the management of physical conditions, fragmented medical care, and patient psychotic beliefs. To improve clinical practice, they suggested reminding clinicians of the latest guidelines through alerts on patient electronic records, having clear policies at the system level, and providing clinicians with training sessions on AF-related stroke management. Clinicians reported many potential benefits for the eCDSS, including improving clinical effectiveness, better identification of patients at risk, safer and more comprehensive care, consistency in decision-making, and saving time. However, they noted that the digital tool could have potential risks such as rigidity in decision-making, overreliance, reduced critical thinking, false positive recommendations, annoyance, and increased workload.

### Comparison to Prior Work

Physical comorbidities among people with mental illness present complex clinical scenarios that require a specialized and holistic approach to care. Fragmentation between primary and secondary health services could contribute to uncertainty regarding which provider is responsible for the management of physical conditions among people with mental illness [21]. This could result in missed opportunities for the identification of physical conditions, which may be hampered by often poor(er) documentation in mental health services [21]. Additionally, inadequate training and lack of physical care skills may reduce mental health care professionals' confidence in managing physical conditions [22]. Continuous training, access to resources, and specialist support are all factors that may influence the level of confidence in dealing with acute conditions considered out of their specialty [22]. Another common scenario that prevents or delays the management of physical conditions among people with mental illness is diagnostic overshadowing, which refers to the misattribution of physical symptoms to mental illness [23]. Features of the mental illness itself may also create major challenges, as people experiencing cognitive impairment, hallucinations, or delusions may not recognize or have difficulty communicating symptoms, may resist medication or struggle with medication adherence [23].

The impact of eCDSSs on AF knowledge, OAC prescription, adherence to guidelines, and patient outcomes has been investigated in general health care settings [10-19], with mixed findings on their effectiveness [10-19]. Research aiming to understand clinician perception of how an eCDSS can be supportive and useful is scarce, although this could serve as a

basis for creating digital health tools that are impactful and aligned with their needs. In a study conducted in China to evaluate the acceptance of an eCDSS that automatically assesses the risks of stroke and bleeding and suggests treatments accordingly, GPs showed positive attitudes toward the digital tool, reporting that it would be helpful and would strengthen their confidence and capabilities in managing patients with AF [24]. This is consistent with results of our study, where clinicians expressed a lack of confidence in managing stroke risk related to AF and their need to refer to guidelines or to seek advice from specialists even if they already knew about current recommendations. Thus, implementing an eCDSS providing the latest guidelines, tools required to complete clinical assessments for stroke and bleeding risks, and guidance on how to interpret these scores would decrease dependence on specialist inputs and increase clinical efficiency. Our findings are also in line with those of a recent systematic review aiming to identify barriers and facilitators of using CDSSs by primary care professionals [25]. In this review, the reported benefits of the digital tools were improving quality of care, saving time, facilitating decision-making, improving professional self-confidence, and updating knowledge [25]. The main barriers were resistance or reluctance, alert fatigue, information overload, disruption of workflow, negative attitude, lack of motivation to use, lack of computer skills, and validity concerns [25].

### Strengths and Limitations

This study has several strengths. First, it used both quantitative and qualitative data collection and analysis methods, which provided a comprehensive and holistic understanding of the topic of interest. While the quantitative methods offered numerical data, the qualitative approach allowed for a deeper exploration of clinician perceptions and experiences with the digital tool. Second, the study was conducted in 3 wards at South London and Maudsley NHS Foundation Trust (SLaM), which enhances the robustness, applicability, and impact of the research findings in a specific health care context. Third, 2 researchers independently worked on data extraction and analysis, which increased the rigor, transparency, and reliability of the research process. This approach also helped reduce researcher bias and validated the results.

However, the study has several limitations. First, there may have been some reluctance among health care professionals to express their lack of knowledge or confidence in assessing physical health conditions, potentially leading to reporting bias. This was mitigated by explicitly informing clinicians that the data from interviews would be deidentified and that the purpose of the study was to understand their experiences in managing AF-related stroke risk and to inform the implementation of an eCDSS in a helpful way. Second, the study focused only on psychiatrists and pharmacists, as they are typically the professionals involved in clinical assessments related to stroke and bleeding risks. Including other health care professionals with diverse clinical experiences might have enriched the findings and provided a broader perspective. Third, the sample size in this study was relatively small. However, this limitation was addressed by continuing data collection until saturation was reached, ensuring that no new themes emerged from participants' perspectives. This approach is consistent with findings from a recent systematic review, which suggests that saturation in qualitative research can typically be achieved within 9-17 interviews [26]. Future studies could consider expanding the sample size to confirm the findings and improve the generalizability of the results.

### Conclusions

The study findings indicate that adoption of an eCDSS for stroke risk screening in a psychiatric health service has the potential to be a valuable tool. However, health care organizations and clinicians need to be mindful of the challenges associated with increased workload and the potential overreliance on the system's recommendations. To maximize the clinical benefits while minimizing the drawbacks, a balanced approach to eCDSS integration is essential. This might involve ongoing training, customization of the system to local practice, and clear guidelines on how to use eCDSS recommendations in conjunction with clinical judgment to provide patient-centered care.

## Acknowledgments

The successful completion of this study was made possible through the gracious cooperation and invaluable contributions of numerous individuals. We would like to express our deepest gratitude to the ward managers who facilitated access to clinicians on their respective wards. We would like to thank all the clinicians who participated in this study. Their willingness to share their insights and experiences played a pivotal role in shaping the findings and outcomes of our research. We are truly grateful for their time, openness, and collaboration. DF, FG, and MA are supported by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust. JO and FG are in part supported by the National Institute for Health Research's (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, and by the Maudsley Charity. For JO, this paper represents independent research funded by the Wellcome Trust [308556/Z/23/Z]. The funders had no involvement in study design, data collection, analysis, interpretation or the decision to submit for publication. The views expressed are those of the author(s) and not necessarily those of the funders. DF is supported by the National Institute for Health Research (NIHR) Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust and by the KCL funded Centre for Doctoral Training (CDT) in Data-Driven Health. MA is in part funded by the Guy's and St Thomas' Charity and the National Institute for Health Research (NIHR) Applied Research Collaboration South London (NIHR ARC South London).

The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

## Data Availability

The datasets used or analyzed in this study are available from the corresponding author on reasonable request.

## Authors' Contributions

FG and MA supervised the study; FG, MA, and DF designed the study; DF collected data; DF and HC analyzed and interpreted data; JO, FG, and MA revised data analysis and interpretation; DF wrote the first manuscript draft. All authors commented on the first draft. All authors read and approved the final manuscript.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Clinician survey.
[DOCX File, 60 KB - cardio_v9i1e66428_app1.docx ]

Multimedia Appendix 2
Semistructured interview topic guide.
[DOCX File, 59 KB - cardio_v9i1e66428_app2.docx ]

## References

1. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. Br Med J 2005 Apr 2;330(7494):765. [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]
2. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3(1):17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
3. Muhiyaddin R, Abd-Alrazaq AA, Househ M, Alam T, Shah Z. The impact of Clinical Decision Support Systems (CDSS) on physicians: a scoping review. Stud Health Technol Inform 2020 Jun 26;272:470-473. [doi: 10.3233/SHTI200597] [Medline: 32604704]
4. Moja L, Kwag KH, Lytras T, et al. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. Am J Public Health 2014 Dec;104(12):e12-e22. [doi: 10.2105/AJPH.2014.302164] [Medline: 25322302]
5. Alasiri SF, Douiri A, Altukistani S, Porat T, Mousa O. The role of clinical decision support systems in preventing stroke in primary care: a systematic review. Perspect Health Inf Manag 2023;20(2):1d. [Medline: 37293480]
6. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Kardiol Pol 2016;74(12):1359-1469. [doi: 10.5603/KP.2016.0172] [Medline: 28009037]
7. Jones C, Pollit V, Fitzmaurice D, Cowan C, Guideline Development Group. The management of atrial fibrillation: summary of updated NICE guidance. Br Med J 2014 Jun 19;348:g3655. [doi: 10.1136/bmj.g3655] [Medline: 24948694]
8. Ahn HJ, Lee SR, Choi EK, et al. Increased risk of incident atrial fibrillation in young adults with mental disorders: a nationwide population-based study. Heart Rhythm 2023 Mar;20(3):365-373. [doi: 10.1016/j.hrthm.2022.12.019] [Medline: 36563829]
9. Farran D, Feely O, Ashworth M, Gaughran F. Anticoagulation therapy and outcomes in patients with atrial fibrillation and serious mental illness: a systematic review and meta-analysis. J Psychiatr Res 2022 Dec;156:737-753. [doi: 10.1016/j.jpsychires.2022.11.002] [Medline: 36417811]
10. Wess ML, Schauer DP, Johnston JA, et al. Application of a decision support tool for anticoagulation in patients with non-valvular atrial fibrillation. J Gen Intern Med 2008 Apr;23(4):411-417. [doi: 10.1007/s11606-007-0477-9] [Medline: 18373138]
11. Cook DA, Enders F, Caraballo PJ, Nishimura RA, Lloyd FJ. An automated clinical alert system for newly-diagnosed atrial fibrillation. PLoS ONE 2015;10(4):e0122153. [doi: 10.1371/journal.pone.0122153] [Medline: 25849969]
12. Robson J, Dostal I, Mathur R, et al. Improving anticoagulation in atrial fibrillation: observational study in three primary care trusts. Br J Gen Pract 2014 May;64(622):e275-e281. [doi: 10.3399/bjgp14X679705] [Medline: 24771841]
13. Arts DL, Abu-Hanna A, Medlock SK, van Weert H. Effectiveness and usage of a decision support system to improve stroke prevention in general practice: a cluster randomized controlled trial. PLoS One 2017;12(2):e0170974. [doi: 10.1371/journal.pone.0170974] [Medline: 28245247]
14. Karlsson LO, Nilsson S, Bång M, Nilsson L, Charitakis E, Janzon M. A clinical decision support tool for improving adherence to guidelines on anticoagulant therapy in patients with atrial fibrillation at risk of stroke: a cluster-randomized

trial in a Swedish primary care setting (the CDS-AF study). PLoS Med 2018 Mar;15(3):e1002528. [doi: 10.1371/journal.pmed.1002528] [Medline: 29534063]
15. Thomson RG, Eccles MP, Steen IN, et al. A patient decision aid to support shared decision-making on anti-thrombotic treatment of patients with atrial fibrillation: randomised controlled trial. Qual Saf Health Care 2007 Jun;16(3):216-223. [doi: 10.1136/qshc.2006.018481] [Medline: 17545350]
16. Fraenkel L, Street RL Jr, Towle V, et al. A pilot randomized controlled trial of a decision support tool to improve the quality of communication and decision-making in individuals with atrial fibrillation. J Am Geriatr Soc 2012 Aug;60(8):1434-1441. [doi: 10.1111/j.1532-5415.2012.04080.x] [Medline: 22861171]
17. Guo Y, Chen Y, Lane DA, Liu L, Wang Y, Lip GYH. Mobile health technology for atrial fibrillation management integrating decision support, education, and patient involvement: mAF app trial. Am J Med 2017 Dec;130(12):1388-1396. [doi: 10.1016/j.amjmed.2017.07.003] [Medline: 28847546]
18. Sheibani R, Sheibani M, Heidari-Bakavoli A, Abu-Hanna A, Eslami S. The effect of a clinical decision support system on improving adherence to guideline in the treatment of atrial fibrillation: an interrupted time series study. J Med Syst 2018 Feb;42(2). [doi: 10.1007/s10916-017-0881-6]
19. Simmons BJ, Jenner KM, Delate T, Clark NP, Kurz D, Witt DM. Pilot study of a novel patient self-management program for warfarin therapy using venipuncture-acquired international normalized ratio monitoring. Pharmacotherapy 2012 Dec;32(12):1078-1084. [doi: 10.1002/phar.1139] [Medline: 23112110]
20. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
21. Stange KC. The problem of fragmentation and the need for integrative solutions. Ann Fam Med 2009;7(2):100-103. [doi: 10.1370/afm.971] [Medline: 19273863]
22. Terry J, Cutter J. Does education improve mental health practitioners' confidence in meeting the physical health needs of mental health service users? A mixed methods pilot study. Issues Ment Health Nurs 2013 Apr;34(4):249-255. [doi: 10.3109/01612840.2012.740768] [Medline: 23566187]
23. Nash M. Diagnostic overshadowing: a potential barrier to physical health care for mental health service users. Mental Health Practice 2013 Dec;17(4):22-26. [doi: 10.7748/mhp2013.12.17.4.22.e862]
24. Ru X, Wang T, Zhu L, et al. Using a clinical decision support system to improve anticoagulation in patients with nonvalve atrial fibrillation in China's primary care settings: a feasibility study. Int J Clin Pract 2023;2023:2136922. [doi: 10.1155/2023/2136922] [Medline: 36713952]
25. Meunier PY, Raynaud C, Guimaraes E, Gueyffier F, Letrilliart L. Barriers and facilitators to the use of clinical decision support systems in primary care: a mixed-methods systematic review. Ann Fam Med 2023;21(1):57-69. [doi: 10.1370/afm.2908] [Medline: 36690490]
26. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523. [doi: 10.1016/j.socscimed.2021.114523]

## Abbreviations

**AF:** atrial fibrillation
**eCDSS:** electronic clinical decision support systems
**EHR:** electronic health record
**GP:** general practitioner
**MHOA:** mental health of older adult
**NHS:** National Health Service
**NICE:** National Institute for Health and Care Excellence
**OAC:** oral anticoagulation
**ORBIT:** Outcomes Registry for Better Informed Treatment of Atrial Fibrillation
**SLaM:** South London and NHS Foundation Trust

*Edited by A Coristine; submitted 12.09.24; peer-reviewed by F Shaikh, J Steinmiller; revised version received 23.12.24; accepted 17.01.25; published 06.08.25.*

*Please cite as:*
*Farran D, Cheang HW, Onwumere J, Ashworth M, Gaughran F*
*Electronic Clinical Decision Support System for Stroke Risk Screening in Patients With Atrial Fibrillation in Mental Health Care: Mixed Methods Study*
*JMIR Cardio 2025;9:e66428*
*URL: https://cardio.jmir.org/2025/1/e66428*
*doi:10.2196/66428*

# Machine Learning Model for Predicting Coronary Heart Disease Risk: Development and Validation Using Insights From a Japanese Population–Based Study

Thien Vu[1,2,3], MD, PhD; Yoshihiro Kokubo[4], MD, PhD; Mai Inoue[1], ME; Masaki Yamamoto[1], BE; Attayeb Mohsen[1,5], MD, PhD; Agustin Martin-Morales[1], PhD; Research Dawadi[1], PhD; Takao Inoue[1,6], PhD; Jie Ting Tay[1], MRES; Mari Yoshizaki[1], PhD; Naoki Watanabe[1], PhD; Yuki Kuriya[1], PhD; Chisa Matsumoto[4,7], MD, PhD; Ahmed Arafa[4,8], MD, PhD; Yoko M Nakao[4], MD, PhD; Yuka Kato[4,9], MD, PhD; Masayuki Teramoto[4], MD, MPH; Michihiro Araki[1,10,11], PhD

[1]Artificial Intelligence Center for Health and Biomedical Research, National Institutes of Biomedical Innovation, Health and Nutrition, 3-17 Senrioka-shinmachi, Osaka, Japan

[10]Graduate School of Medicine, Kyoto University, Kyoto, Japan

[11]Graduate School of Science, Technology and Innovation, Kobe University, Kobe, Japan

[2]NCD Epidemiology Research Center, Shiga University of Medical Science, Shiga, Otsu, Japan

[3]Department of Cardiac Surgery, Cardiovascular Center, Cho Ray Hospital, Ho Chi Minh City, Vietnam

[4]Department of Preventive Cardiology, National Cerebral and Cardiovascular Center, Suita, Osaka, Japan

[5]Libyan Centre for Dental Research, Zliten, Libya

[6]Faculty of Informatics, Yamato University, Osaka, Japan

[7]Department of Cardiology, Center for Health Surveillance and Preventive Medicine, Tokyo Medical University Hospital, Tokyo, Japan

[8]Department of Public Health, Faculty of Medicine, Beni-Suef University, Beni-Suef, Egypt

[9]Division of Health Sciences, Osaka University Graduate School of Medicine, Suita, Osaka, Japan

**Corresponding Author:**
Thien Vu, MD, PhD
Artificial Intelligence Center for Health and Biomedical Research, National Institutes of Biomedical Innovation, Health and Nutrition, 3-17 Senrioka-shinmachi, Osaka, Japan

## Abstract

**Background:** Coronary heart disease (CHD) is a major cause of morbidity and mortality worldwide. Identifying key risk factors is essential for effective risk assessment and prevention. A data-driven approach using machine learning (ML) offers advanced techniques to analyze complex, nonlinear, and high-dimensional datasets, uncovering novel predictors of CHD that go beyond the limitations of traditional models, which rely on predefined variables.

**Objective:** This study aims to evaluate the contribution of various risk factors to CHD, focusing on both established and novel markers using ML techniques.

**Methods:** The study recruited 7672 participants aged 30-84 years from Suita City, Japan, between 1989 and 1999. Over an average of 15 years, participants were monitored for cardiovascular events. A total of 7260 participants and 28 variables were included in the analysis after excluding individuals with missing outcome data and eliminating unnecessary variables. Five ML models—logistic regression, random forest (RF), support vector machine, Extreme Gradient Boosting, and Light Gradient-Boosting Machine—were applied for predicting CHD incidence. Model performance was evaluated using accuracy, sensitivity, specificity, precision, area under the curve, $F_1$-score, calibration curves, observed-to-expected ratios, and decision curve analysis. Additionally, Shapley Additive Explanations (SHAPs) were used to interpret the prediction models and understand the contribution of various risk factors to CHD.

**Results:** Among 7260 participants, 305 (4.2%) were diagnosed with CHD. The RF model demonstrated the highest performance, with an accuracy of 0.73 (95% CI 0.64 - 0.80), sensitivity of 0.74 (95% CI 0.62 - 0.84), specificity of 0.72 (95% CI 0.61 - 0.83), and an area under the curve of 0.73 (95% CI 0.65 - 0.80). RF also showed excellent calibration, with predicted probabilities closely aligning with observed outcomes, and provided substantial net benefit across a range of risk thresholds, as demonstrated by decision curve analysis. SHAP analysis elucidated key predictors of CHD, including the intima-media thickness (IMT_cMax) of the common carotid artery, blood pressure, lipid profiles (non–high-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and triglycerides), and estimated glomerular filtration rate. Novel risk factors identified as significant contributors to CHD risk included lower calcium levels, elevated white blood cell counts, and body fat percentage. Furthermore, a protective

effect was observed in women, suggesting the potential necessity for gender-specific risk assessment strategies in future cardiovascular health evaluations.

**Conclusions:** We developed a model to predict CHD using ML and applied SHAP methods for interpretation. This approach highlights the multifactor nature of CHD risk evaluation, aiming to support health care professionals in identifying risk factors and formulating effective prevention strategies.

***(JMIR Cardio 2025;9:e68066)*** doi:10.2196/68066

**KEYWORDS**

coronary heart disease; machine learning; logistic regression; random forest; support vector machine; Extreme Gradient Boosting; Light Gradient-Boosting Machine; Shapley Additive Explanations; CHD; SVM; XGBoost; LightGBM; SHAP

## Introduction

Coronary heart disease (CHD) remains a leading cause of morbidity and mortality worldwide, responsible for approximately 9.14 million deaths in 2019 [1,2]. Early identification of individuals at high risk is crucial, as timely interventions can significantly reduce the likelihood of severe outcomes like heart attacks and strokes. Studies have shown that early prediction and intervention can lead to a notable reduction in CHD-related mortality through preventive treatments such as statins and lifestyle changes [1-3]. While conventional risk assessment models have been used, there is growing recognition of the potential of machine learning (ML) in enhancing CHD prediction [4,5].

ML algorithms have proven their ability to analyze complex data and identify intricate patterns and relationships that are not easily detected by traditional statistical methods [6-10]. By integrating diverse data sources, such as demographics, medical history, lifestyle habits, and diagnostic findings, these algorithms can predict the likelihood of developing CHD. This approach offers comprehensive risk evaluation, adaptability to new data, and the potential to uncover novel risk factors and disease mechanisms [11].

Several studies have demonstrated the effectiveness of ML models in deriving quantitative markers for coronary artery disease and predicting the presence of heart disease. For example, a study developed and validated a coronary artery disease–predictive ML model using electronic health records and assessed its probabilities as in silico scores for coronary artery disease in participants in 2 longitudinal biobank cohorts [12]. Another study applied an ensemble ML model for coronary disease prediction, using ML classifiers to predict heart disease [13]. These findings highlight the potential of ML in driving innovation and improving the accuracy of CHD diagnosis and prediction [14].

However, challenges exist in utilizing ML for CHD prediction, including data quality, feature selection, model interpretability, and generalizability. These issues must be carefully addressed to ensure the reliability and robustness of the predictive models. Rigorous validation, regulatory compliance, and effective communication strategies are essential for its successful integration into clinical practice.

While several established CHD prediction models rely on traditional statistical techniques with predefined risk factors, they are limited by linear assumptions and struggle with complex, high-dimensional datasets. This restricts their ability to uncover novel or subtle risk factors. In contrast, ML models can handle these complexities, offering more nuanced and accurate predictions by identifying nonlinear interactions and discovering previously overlooked factors. Therefore, ML may enhance the overall understanding of CHD and improve both risk assessment and prevention strategies.

This study aimed to address the role of ML techniques in predicting incident CHD and identifying novel risk factors. This study sought to deepen our understanding of the factors contributing to CHD development by analyzing a comprehensive dataset. These findings will enhance risk assessment, enabling the development of personalized interventions and preventive strategies.

## Methods

### Study Design and Participants

The Suita Study, a prospective population-based cohort study, was conducted in Suita City, Osaka, Japan. From 1989 to 1999, a total of 7672 men and women aged 30-84 years who did not have a previous history of cardiovascular disease were recruited for the study. Participants were selected from the population registry of the municipality and were followed up every 2 years for an average of 15 years until their first occurrence of stroke, myocardial infarction (MI), death, or relocation.

After excluding participants with missing outcome data and removing unnecessary variables, the analysis included 7260 participants and 28 variables. Opt-out procedures were implemented for those who preferred not to participate in this study. Informed consent was obtained from all participants at the time of enrollment. The study followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis and Artificial Intelligence (TRIPOD+AI) statement guidelines for reporting prediction models in medicine, and we have added the completed checklist in Checklist 1 [15].

### Ethical Considerations

The study was conducted in compliance with the ethical standards outlined in the Declaration of Helsinki, and approval was granted by the Institutional Review Board at the National Cerebral and Cardiovascular Center (approval R21024-2). As this study involves secondary data analysis, it is important to note that the original informed consent, obtained during the primary data collection, permits the use of the data for secondary

analyses without requiring additional consent from participants. Participants' privacy was protected by anonymizing or deidentifying the data to prevent identification.

## Outcome

The primary outcome was CHD, including MI, sudden death within 24 hours of acute illness onset, and coronary artery disease requiring bypass surgery or intervention. Medical records were carefully reviewed by hospital doctors or researchers who were blinded to the baseline data to provide an unbiased approach to the analysis. MIs were classified as definite or probable according to the criteria established by the MONICA Project [16].

Every 2 years, each participant's health was evaluated at the National Cerebral and Cardiovascular Center in Osaka, Japan, to detect the occurrence of CHD. Yearly questionnaires were also completed by all participants by mail or telephone. CHD surveillance was completed by systematically searching for death certificates [17,18].

## Predictors

Predictors were measured at baseline and processed according to a standardized protocol. A comprehensive and prospective data collection process was implemented, encompassing various aspects such as demographics, medical history, medical imaging, laboratory data, lifestyle habits, and outcomes.

### *Blood Pressure and Physical Measurements*

Blood pressure was measured in each participant using a mercury column sphygmomanometer, an appropriately sized cuff, and a standardized protocol to ensure accuracy and precision [17]. Before the initial blood pressure reading, the participants were instructed to rest for at least 5 minutes to establish a stable baseline. Blood pressure readings were obtained by averaging the second and third measurements, which were performed at intervals of more than 1 minute to allow for adequate observation and recording. Hypertension was defined as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or the use of antihypertensive medications.

BMI was calculated as weight (kg) divided by the square of height ($m^2$).

### *Biochemical Measurements*

At baseline, routine blood tests were conducted, including measurements of total cholesterol, high-density lipoprotein cholesterol (HDL-c), and fasting glucose levels. Non-HDL-c was calculated by subtracting HDL-c from total cholesterol. Diabetes mellitus was diagnosed if participants had fasting plasma glucose (FPG) ≥126 mg/dL, a non-FPG ≥200 mg/dL, or the use of diabetes mellitus medication.

The estimated glomerular filtration rate (eGFR; mL/min/1.73 $m^2$) was calculated according to the original Modification of Diet in Renal Disease equation modified by the Japanese coefficient (0.881) as follows: $0.881 \times 186 \times \text{serum creatinine}^{-1.154} \times \text{age}^{-0.203} \times (0.742 \text{ if female})$ [19].

### *Imaging Diagnostics*

Carotid artery measurements were performed using a high-resolution ultrasound machine to assess atherosclerotic indices, specifically intima-media thickness (IMT), on both sides of the common carotid artery (CCA), carotid artery bulb, internal carotid artery, and external carotid artery. The maximum IMT in the CCA (IMT_cMax) was defined as the highest measurable IMT in the scanned CCA regions, while the maximum IMT (IMT_MAX) was the highest measurable IMT across the entire scanned area, including the CCA, bulb, internal carotid artery, and external carotid artery on both sides [20].

Atrial fibrillation was checked by standard 12-lead ECGs from all participants and was determined by well-trained physicians [18].

### *Lifestyle and Medical History*

Smoking status and drinking statuses were categorized as current, quit, or never. A questionnaire was used to ask participants about their past and present history of CHD.

## Sample Size

All available data were used, and no formal sample size calculation was performed. The dataset included 7260 participants, among whom 305 had CHD, with 28 predictors selected after the feature selection process used in the model. Based on the events per predictor ratio, which is approximately 10.89 (305/28), the sample size is sufficient to ensure model stability and reliability [21,22]. Therefore, this dataset is adequate to answer the research questions.

## Missing Data

Missing data analysis was conducted, and variables with more than 30% missing values were excluded to enhance model robustness. Missing data were imputed using Multivariate Imputation by Chained Equations. See Multimedia Appendix 1 for details on the percentage of missing data for each variable before imputation.

## Statistical Analysis Methods

### *Descriptive Analysis*

Continuous variables were summarized using means and SDs for normally distributed data, or medians and IQR for nonnormally distributed data. Categorical variables were reported as frequencies and percentages. To compare differences in patient characteristics based on CHD incidence (yes or no), we used various statistical tests including 2-tailed Student *t* tests, Mann Whitney *U* tests, or chi-square tests, as appropriate.

### *Feature Selection*

Feature selection was conducted in a stepwise manner to ensure that only the most relevant variables were included in the predictive models. Initially, variables with more than 30% missing data were excluded to avoid potential bias from imputation. Following this, a correlation matrix was used to identify and remove variables with high multicollinearity, defined as having a correlation coefficient greater than 0.8. See correlation coefficients heat map in the Multimedia Appendix 2 for details. The next step involved applying the least absolute

shrinkage and selection operator regression. This technique shrinks the coefficients of less significant predictors toward zero, effectively removing them from the model, and was performed using cross-validation to identify the most important features based on the data. Finally, after statistical feature selection, medical knowledge was applied to confirm the clinical relevance of the remaining variables. Important predictors such as age, glucose levels, HDL-c, and blood pressure were retained, given their established association with CHD. The list of variables (predictors) used for model development was described in Multimedia Appendix 3.

### Development of ML Models

#### *Overview*

The goal of this analysis was to predict the incidence of CHD using ML models and examine the contribution of each risk factor to the CHD incidence. A comprehensive process was followed, which included descriptive analysis, feature selection, model training, hyperparameter optimization, and interpretability through Shapley Additive Explanation (SHAP) values.

To manage the imbalance between CHD and non-CHD cases, we used down sampling on the majority class (non-CHD) to create a balanced dataset. This approach helps to ensure that the models do not disproportionately favor the majority class during training, improving prediction performance on the minority class.

The dataset was split into training (80%) and testing (20%) sets while maintaining balanced target variable distributions across both. Next, one-hot encoding was applied to convert categorical variables into a binary format, and normalization was performed to scale numerical features.

Several ML algorithms were implemented to compare their predictive power. Logistic regression (LR) was used as a baseline model, offering simplicity and interpretability [23]. Random forest (RF), an ensemble learning method, was used due to its strength in handling high-dimensional data and offering feature-importance insights [8,24]. Support vector machines (SVMs) with radial basis kernels were used for their effectiveness in nonlinear classification tasks [25,26]. Extreme Gradient Boosting (XGBoost) is an ML algorithm that improves model performance by using a series of decision trees, where each tree corrects the mistakes of the previous one. This sequential approach helps make predictions more accurate. Light Gradient-Boosting Machine (LightGBM) is another efficient algorithm that works similarly to XGBoost but is designed to be faster and more scalable, especially when working with large datasets and many features. Both algorithms are known for their high performance in handling complex data and large-scale problems [9,27].

#### *Model Evaluation*

We used 5-fold cross-validation during model training to ensure robustness and mitigate overfitting. Hyperparameter optimization was conducted using a grid search approach. The model's performance on the testing set was evaluated using 5 metrics: accuracy, sensitivity, specificity, precision, area under the curve (AUC), and $F_1$-score [15].

Calibration plots are used to evaluate the predictive accuracy of ML models in estimating CHD incidence. Calibration measures how closely the predicted absolute risk corresponds to the observed (true) risk across groups of patients categorized into different risk levels. The overall observed-to-expected (OE) ratio is calculated by dividing the total observed events by the total expected events for the entire population. For each decile, the OE ratio is determined by dividing the observed events within that decile by the expected events for the same decile. An ideal model is represented by a straight line bisecting the calibration plot, with an OE ratio of 1, indicating perfect calibration. An OE ratio <1 suggests overprediction, while a ratio >1 indicates underprediction [15].

Decision curve analysis (DCA) assesses the clinical use of ML models for predicting CHD incidence. DCA uses net benefit as a metric, reflecting the tradeoff between true-positive and false-positive predictions for a specific strategy [15,28,29].

#### *Model Interpretation*

SHAP is a method used in ML to make the predictions of a model more understandable. It helps explain how each input feature (such as age, cholesterol levels, or blood pressure) affects the model's decision. Essentially, SHAP breaks down the prediction to show how much each feature contributes to the final result, allowing us to see which factors are most important for predicting a condition like CHD [8-10,30]. SHAP summary plots visualized the importance of key features, while SHAP dependence plots highlighted the non-linear relationships between features and CHD incidence.

## Results

### Study Participants' Characteristics

In this study, 7260 participants were analyzed, of which 305 (4.2%) were diagnosed with CHD. The median age of participants with CHD was 63 (IQR 56-71) years , which was significantly older than that of those without CHD, whose median age was 55 (IQR 44-65) years. CHD was more prevalent in men (n=202, 66.2%) compared to women (n=103, 33.8%), and this gender difference was statistically significant.

Several cardiovascular risk factors were also associated with CHD. Participants with CHD had higher systolic and diastolic blood pressures. The eGFR was lower in participants with CHD compared to those without. The IMT of CCAs, IMT_cMax, was also significantly higher in patients with CHD (1.10 mm vs 1.00 mm; $P$<.001).

BMI and waist circumference were also higher in participants with CHD, indicating a greater degree of obesity. Additionally, lipid profiles showed significant differences, with lower HDL-c levels and higher non-HDL-c and triglyceride levels in patients with CHD.

Higher glucose levels and white blood cell counts were observed in participants with CHD, along with elevated hemoglobin levels. Regarding lifestyle factors, smoking was more common in those with CHD, while drinking status did not differ significantly between the 2 groups.

Regarding lifestyle factors, current smoking was more prevalent among participants with CHD (36.1% vs 29.0%; $P<.001$), while drinking status did not significantly differ between the groups. In terms of comorbidities, atrial fibrillation, hypertension, diabetes mellitus, and dyslipidemia were all significantly more common in participants with CHD, as outlined in Table 1.

**Table .** Characteristics of study participants with and without CHD[a] incidence (Japanese participants aged 30 - 84 years, Suita Study). CHD was diagnosed by a first-ever acute myocardial infarction, sudden cardiac death within 24 hours of illness, or coronary artery disease followed by bypass or angioplasty. Values are presented as mean (SD) for continuous variables with approximately normally distribution or by median (IQR) with skewed distribution and n (%) for categorical variables. Differences in characteristics were evaluated by using the unpaired 2-tailed Student *t* test, Wilcoxon rank sums test, or chi-square test.

| | CHD | | *P* value |
|---|---|---|---|
| | No (n=6955) | Yes (n=305) | |
| Age (years), median (IQR) | 55.0 (44.0-65.0) | 63.0 (56.0-71.0) | <.001 |
| Sex, n (%) | | | <.001 |
| Male | 3147 (45.2) | 202 (66.2) | |
| Female | 3808 (54.8) | 103 (33.8) | |
| SBP[b] (mm Hg), median (IQR) | 123 (110-137) | 138 (125-153) | <.001 |
| DBP[c] (mm Hg), median (IQR) | 77.0 (70.0-85.0) | 83.0 (74.0-89.0) | <.001 |
| IMT_cMax[d] (mm), median (IQR) | 1.00 (0.80-1.10) | 1.10 (1.00-1.30) | <.001 |
| eGFR[e] (mL/min/1.73 m²), mean (SD) | 104 (32.2) | 95.3 (63.3) | .014 |
| BMI (kg/m²), mean (SD) | 22.5 (3.10) | 23.3 (3.26) | <.001 |
| Body fat (%), mean (SD) | 23.2 (6.32) | 22.6 (7.06) | .15 |
| Waist circumference (cm), median (IQR) | 80.0 (73.0-86.0) | 83.0 (77.0-90.0) | <.001 |
| HDL-c[f] (mg/dL), median (IQR) | 53.0 (44.0-63.0) | 46.0 (38.0-56.0) | <.001 |
| non-HDL-c (mg/dL), mean (SD) | 152 (36.9) | 172 (40.5) | <.001 |
| Triglycerides (mg/dL), median (IQR) | 98.0 (70.0-143) | 121 (90.0-174) | <.001 |
| Calcium (mg/dL), mean (SD) | 9.35 (0.46) | 9.34 (0.43) | .61 |
| Fructosamine (µmol/L), median (IQR) | 251 (237-266) | 257 (242-276) | <.001 |
| Glucose (mg/dL), median (IQR) | 95.0 (89.0-101) | 100 (93.0-109) | <.001 |
| WBC[g] count (/mm³), median (IQR) | 5.33 (4.48-6.36) | 5.65 (4.81-6.78) | <.001 |
| RBC[h] count (10³/mm³), mean (SD) | 4.53 (0.44) | 4.60 (0.46) | .008 |
| Smoking status, n (%) | | | <.001 |
| Current | 2019 (29) | 110 (36.1) | |
| Past | 1091 (15.7) | 79 (25.9) | |
| Never | 3845 (55.3) | 116 (38) | |
| Drinking status, n (%) | | | .27 |
| Current | 3613 (51.9) | 152 (49.8) | |
| Past | 156 (2.24) | 11 (3.61) | |
| Never | 3186 (45.8) | 142 (46.6) | |
| Atrial fibrillation, n (%) | 123 (1.77) | 20 (6.56) | <.001 |
| Hypertension, n (%) | 2056 (29.6) | 172 (56.4) | <.001 |
| Diabetes mellitus, n (%) | 426 (6.13) | 49 (16.1) | <.001 |
| Dyslipidemia, n (%) | 5280 (75.9) | 265 (86.9) | <.001 |

[a]CHD: coronary heart disease.

[b]SBP: systolic blood pressure.

[c]DBP: diastolic blood pressure.

[d]IMT_cMax: maximum intima-media thickness of common carotid arteries.

[e]eGFR: estimated glomerular filtration rate.

[f]HDL-c: high-density lipoprotein cholesterol.

[g]WBC: white blood cell.

[h]RBC: red blood cell.

## Model Performance

The performance metrics of the 5 ML models used in our CHD prediction study provide valuable insights into their effectiveness, as shown in Table 2.

**Table .** Performance metrics and 95% CIs for machine learning models predicting CHD[a] incidence (Japanese participants, aged 30 - 84 years, Suita Study).

| Model | Accuracy | Sensitivity | Specificity | Precision | AUC[b] | $F_1$-score |
|---|---|---|---|---|---|---|
| LR[c] | 0.66 (0.58 - 0.75) | 0.59 (0.46 - 0.71) | 0.74 (0.62 - 0.84) | 0.69 (0.55 - 0.81) | 0.66 (0.57 - 0.75) | 0.64 (0.52 - 0.73) |
| RF[d] | 0.73 (0.64 - 0.80) | 0.74 (0.62 - 0.84) | 0.72 (0.61 - 0.83) | 0.73 (0.61 - 0.84) | 0.73 (0.65 - 0.80) | 0.73 (0.64 - 0.82) |
| SVM[e] | 0.71 (0.62 - 0.80) | 0.70 (0.59 - 0.81) | 0.72 (0.62 - 0.83) | 0.72 (0.60 - 0.84) | 0.71 (0.63 - 0.79) | 0.71 (0.61 - 0.80) |
| XGBoost[f] | 0.72 (0.64 - 0.80) | 0.74 (0.63 - 0.84) | 0.70 (0.58 - 0.82) | 0.71 (0.60 - 0.82) | 0.72 (0.64 - 0.80) | 0.73 (0.63 - 0.81) |
| LightGBM[g] | 0.50 (0.43 - 0.58) | 1.00 (1.00 - 1.00) | 0.00 (0.00 - 0.00) | 0.50 (0.41 - 0.59) | 0.5 (0.49 - 0.57) | 0.67 (0.58 - 0.74) |

[a]CHD: coronary heart disease.

[b]AUC: area under the curve.

[c]LR: logistic regression.

[d]RF: random forest.

[e]SVM: support vector machine.

[f]XGBoost: Extreme Gradient Boosting.

[g]LightGBM: Light Gradient-Boosting Machine.

RF emerged as the strongest model for CHD prediction in this study, achieving the highest overall performance with an accuracy of 0.73 (95% CI 0.64 - 0.80), sensitivity of 0.74 (95% CI 0.62 - 0.84), specificity of 0.72 (95% CI 0.61 - 0.83), and an AUC of 0.73 (95% CI 0.65 - 0.80). These results highlight its balanced ability to identify both CHD and non-CHD cases effectively. In comparison, XGBoost delivered robust, yet slightly inferior, results with an accuracy of 0.72 (95% CI 0.64 - 0.80), sensitivity of 0.74 (95% CI 0.63 - 0.84), specificity of 0.70 (95% CI 0.58 - 0.82), an AUC of 0.72 (95% CI 0.64 - 0.80), and an $F_1$-score of 0.73 (95% CI 0.63 - 0.81). SVM demonstrated competitive performance, achieving an AUC of 0.71 (95% CI 0.63 - 0.79), but ranked slightly behind RF and XGBoost. In contrast, LightGBM, despite its perfect sensitivity of 1.00 (95% CI 1.00 - 1.00), showed a specificity of 0.00 (95% CI 0.00 - 0.00) and an AUC of 0.50 (95% CI 0.49 - 0.57), rendering it unsuitable for this task. LR, while serving as a baseline model, exhibited moderate performance with an accuracy of 0.66 (95% CI 0.58 - 0.75), sensitivity of 0.59 (95% CI 0.46 - 0.71), specificity of 0.74 (95% CI 0.62 - 0.84), and an AUC of 0.66 (95% CI 0.57 - 0.75), but lacked the sensitivity required for effective CHD prediction.

The calibration curves for the 5 models (Figure 1) and the OE ratios by decile (Figure 2) provide critical insights into their predictive reliability. Among the models, RF demonstrated excellent calibration, with predicted probabilities closely aligning with observed outcomes across all deciles. This strong calibration is complemented by its performance in DCA (Figure 3).

**Figure 1.** Calibration plots for machine learning models predicting CHD incidence (Japanese participants, aged 30 - 84 years, Suita Study). CHD: coronary heart disease; LightGBM: Light Gradient-Boosting Machine; XGBoost: Extreme Gradient Boosting.

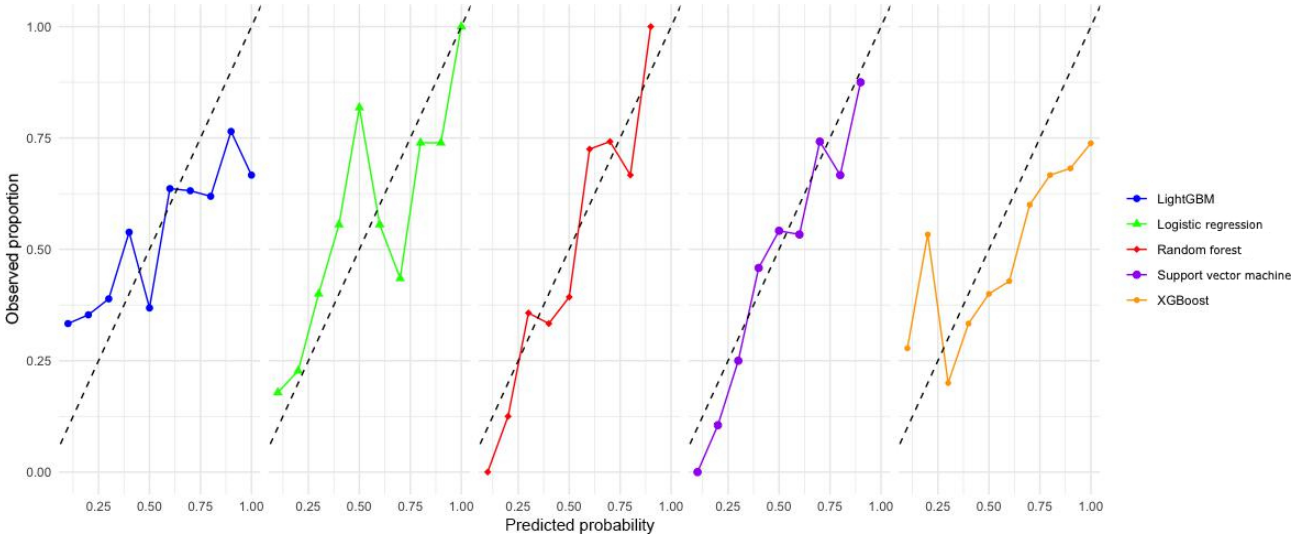



**Figure 2.** Calibration plots displaying observed-to-expected ratios for each decile of predicted CHD incidence risk (Japanese participants, aged 30 - 84 years, Suita Study). CHD: coronary heart disease; LightGBM: Light Gradient-Boosting Machine; SVM: support vector machine; XGBoost: Extreme Gradient Boosting.

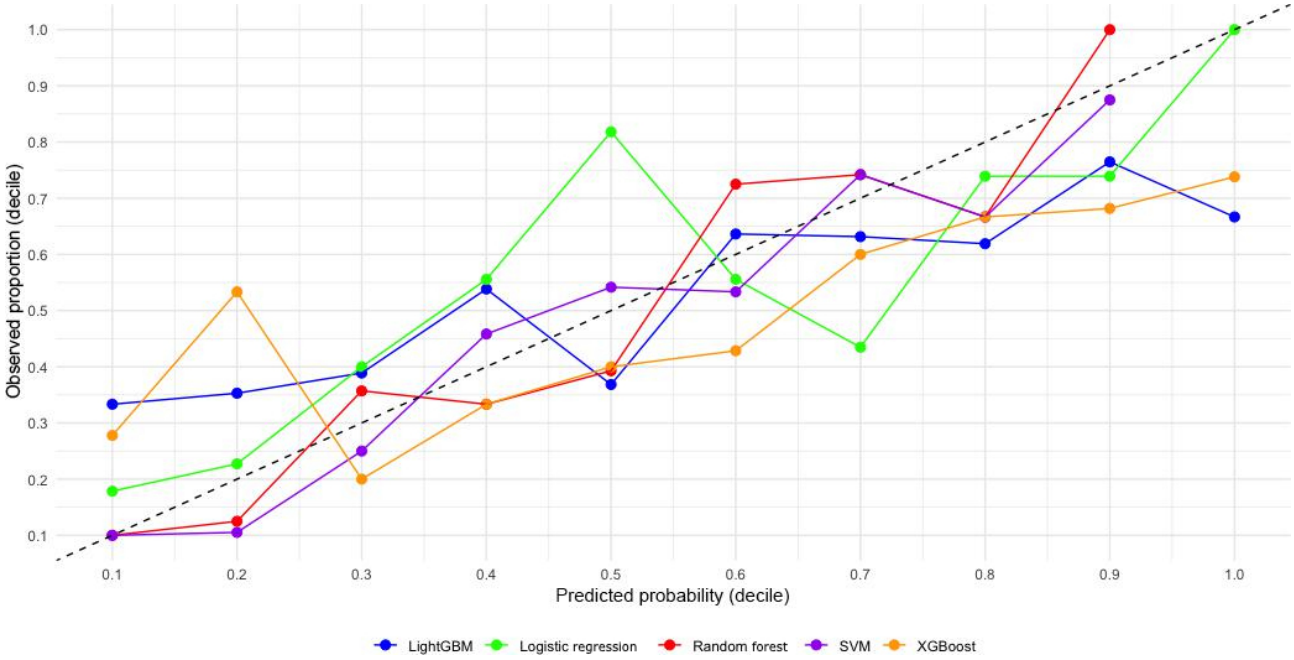
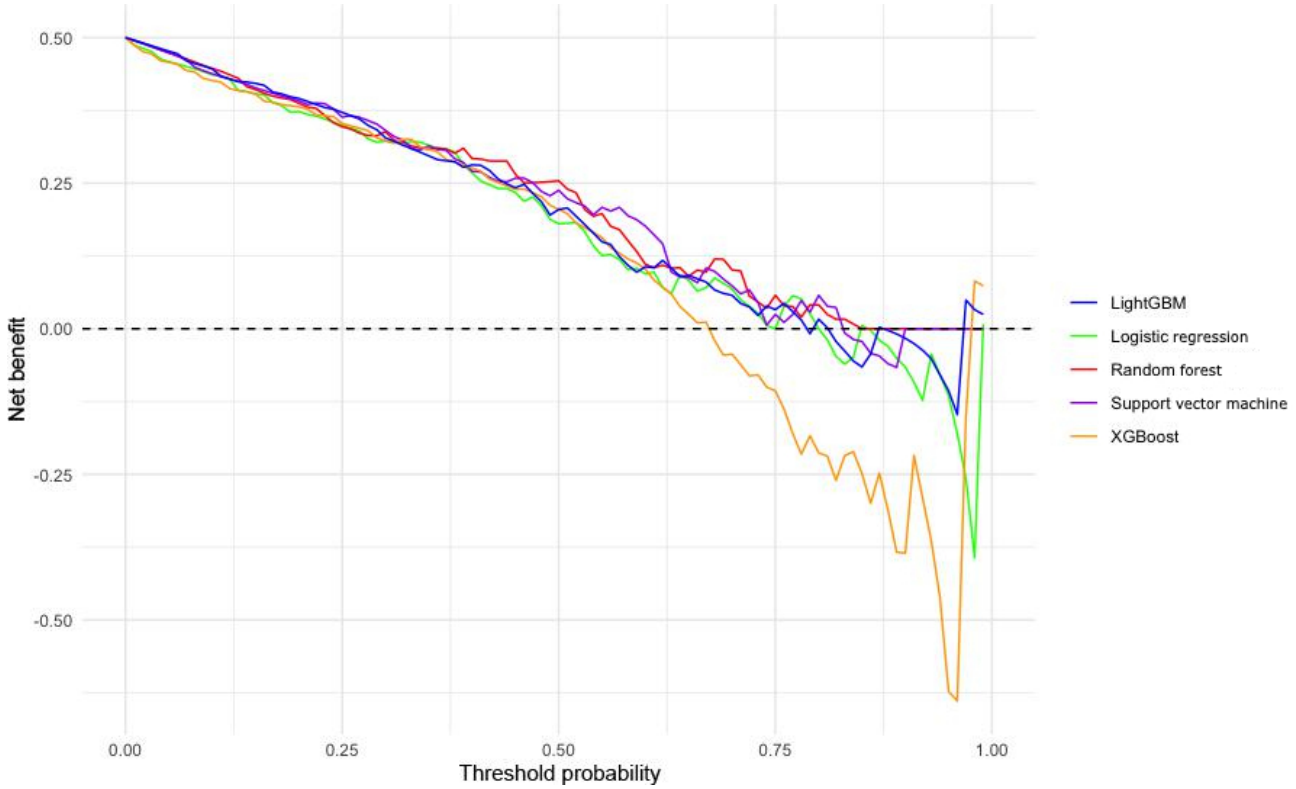


**Figure 3.** Decision curve analysis comparing machine learning models for predicting CHD incidence (Japanese participants, aged 30 - 84 years, Suita Study). CHD: coronary heart disease; LightGBM: Light Gradient-Boosting Machine; XGBoost: Extreme Gradient Boosting.




In terms of clinical use, as illustrated in Figure 3, all models exhibit a similar positive net benefit when the threshold is below 0.5, meaning that using the predictive models is better than not using any model (treat none). However, when the threshold exceeds 0.5, the models tend to decline rapidly, with LR and XGBoost showing the most pronounced decrease, declining earlier compared to the other models.

Based on the performance metrics, RF emerges as the best model for CHD prediction in this study due to its highest overall accuracy, balanced sensitivity and specificity, strong AUC, excellent calibration, and robust clinical use across various threshold probabilities.

## Model Interpretation

In Figure 4, the bar plot on the left ranks the top features contributing to CHD prediction, with IMT_cMax identified as the most influential variable, followed by systolic blood pressure (SBP), HDL-c, non-HDL-c, and eGFR. This ranking emphasizes the significance of arterial health, blood pressure regulation, lipid levels, and kidney function in assessing CHD risk. The SHAP summary heat plot on the right provides a detailed visualization of how each feature influences individual model predictions. It shows that higher values of IMT_cMax, non-HDL-c, and blood pressure are positively associated with an increased likelihood of CHD, whereas lower levels of protective factors like HDL-c and eGFR are associated with a higher risk of CHD. Other important variables, such as age, glucose levels, and triglycerides, also contribute significantly, with older age and impaired glucose regulation being linked to a higher CHD risk. Additionally, markers of inflammation like white blood cell count and other factors such as calcium levels, sex, body fat, and BMI play roles in determining CHD risk.

Figure 5 consists of several SHAP dependency plots that illustrate the relationship between each key variable and CHD risk in more detail. For IMT_cMax, there is a positive association with CHD risk, showing that as the thickness of the carotid artery increases, so does the risk of CHD. The eGFR plot shows that lower eGFR values are associated with a higher risk of CHD, while higher eGFR values are associated with a lower risk, indicating the crucial role of kidney function in cardiovascular health. Non-HDL-c shows a generally positive association with CHD, where higher levels correspond to a higher risk. For SBP, the risk of CHD increases sharply with rising SBP values. HDL-c is inversely related to CHD risk, indicating its protective role, while higher triglycerides (TG) are linked to increased risk, especially at moderate levels. Age and glucose levels show a direct relationship with CHD risk, whereas older age and higher glucose levels are associated with increased risk. The SHAP value for diastolic blood pressure (DBP) also shows a positive relationship, suggesting that higher DBP levels contribute to the increased risk of CHD.

**Figure 4.** Contribution of variables to CHD incidence prediction using SHAP values (Japanese participants, aged 30 - 84 years, Suita Study). (A) The bar plot shows each variable's contribution to CHD, with bar length indicating the contribution extent. (B) The heat plot of SHAP values illustrates the relationships between variables and CHD. Purple signifies a positive relationship and yellow a negative one. Each point represents a participant, with the x-axis showing SHAP values and the y-axis indicating variable importance. bf: body fat; Ca: calcium; CHD: coronary heart disease; DBP: diastolic blood pressure; eGFR: estimated glomerular filtration rate; Frct: Fructosamine; Hb: hemoglobin; htn: hypertension; IMT_cMax: maximum intima-media thickness of common carotid arteries; HDL-c: high-density lipoprotein cholesterol; SBP: systolic blood pressure; smk_sts: smoking status; TG: triglycerides; WBC: white blood cell; wt20: weight at age of 20 years.

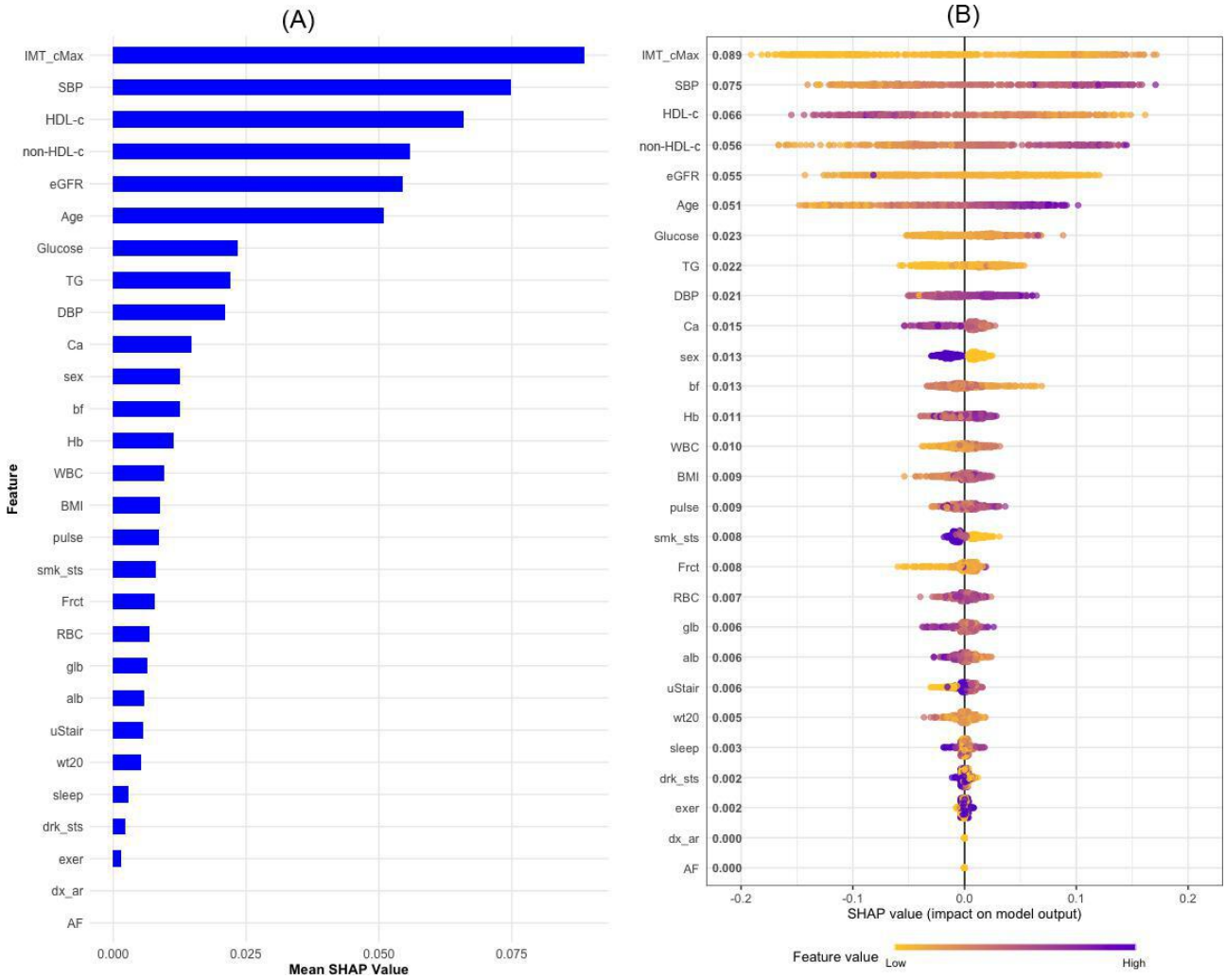
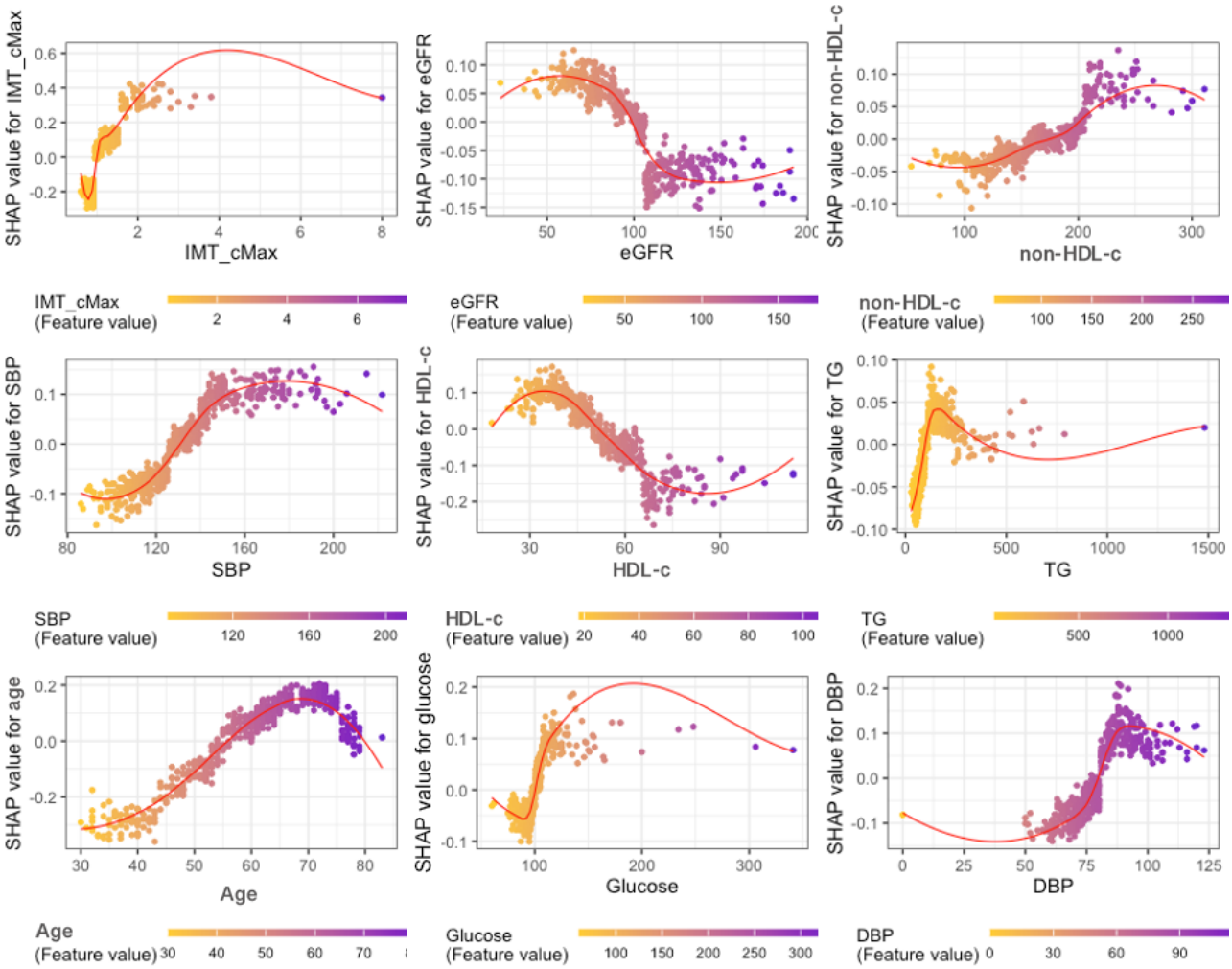


**Figure 5.** SHAP dependency plots showing the relationship between key variables and CHD risk (Japanese participants, aged 30 - 84 years, Suita Study). CHD: coronary heart disease; DBP: diastolic blood pressure; eGFR: estimated glomerular filtration rate; IMT_cMax: maximum intima-media thickness of common carotid arteries; HDL-c: high-density lipoprotein cholesterol; SBP: systolic blood pressure; SHAP: Shapley Additive Explanation; TG: triglycerides.




## Discussion

### Principal Findings

This study provides a comprehensive evaluation of the role of ML in predicting CHD. Among a cohort of 7260 participants, 305 were diagnosed with CHD. The analysis not only validated several well-established cardiovascular and metabolic risk factors but also identified novel predictors of CHD. Importantly, the findings underscore the use of ML models and the SHAP method in elucidating key contributors to CHD risk, with RF demonstrating superior performance, excelling in both discrimination and calibration for CHD prediction.

### Comparison With Prior Work

#### *Arterial Health*

Carotid IMT emerged as the strongest predictor of CHD in our study. IMT_cMax, which measures the thickness of the CCAs, is a well-established indicator of atherosclerosis and future cardiovascular events, including MI and stroke [31,32]. Multiple studies support this, showing that even a small increase in IMT correlates with a significantly elevated risk of acute MI and stroke. For instance, in the Atherosclerosis Risk in Communities study, a 0.1 mm increase in IMT corresponded to a 50% increase in CHD risk [20,31]. Therefore, measuring IMT through noninvasive techniques like ultrasound has important clinical applications in evaluating subclinical atherosclerosis and assessing CHD risk. Given that many coronary artery assessments are invasive, the use of ultrasound to measure carotid artery IMT offers a valuable alternative for early detection and risk stratification.

#### *Blood Pressure, Lipid Profiles, and Glucose*

SBP and hypertension were among the most critical predictors of CHD, aligning with the well-established association between elevated blood pressure and cardiovascular risk [33,34]. Both SBP and diastolic blood pressure were prominent, emphasizing the need for effective blood pressure management in reducing CHD risk [33,35].

Furthermore, non-HDL-c and triglycerides were strongly associated with CHD, confirming the importance of lipid management in cardiovascular health [36-39]. Glucose levels were also significant, suggesting that monitoring glucose metabolism is essential in cardiovascular risk management [40-42].

### *Renal Function and Metabolic Factors*

The role of eGFR as a key predictor highlights the connection between renal function and CHD [43]. Impaired kidney function has been increasingly recognized as a cardiovascular risk factor, particularly due to its association with hypertension and dyslipidemia [44,45]. The results support incorporating kidney function markers in future CHD risk assessments. In addition, metabolic markers and body fat percentage were identified as important predictors, signaling the impact of obesity-related factors on cardiovascular health. These findings suggest that obesity-related measures beyond BMI should be considered in CHD risk assessments.

### *Sex*

The sex-specific analysis highlighted the protective effect of being female, consistent with existing research showing that premenopausal women are generally at a lower risk of developing CHD due to protective hormonal factors [46,47]. These findings suggest the need for sex-specific strategies in managing CHD risk.

## Potential Risk Factors

One of the strengths of this study is its ability to uncover novel predictors, such as white blood cell count, which serves as a marker of systemic inflammation. Inflammation is increasingly recognized as a key player in the development of atherosclerosis and cardiovascular events. Additionally, lower calcium levels were associated with a higher risk of CHD, highlighting the importance of mineral balance in cardiovascular health. Furthermore, body fat percentage and BMI were highlighted as significant predictors of CHD, further emphasizing the need for a comprehensive evaluation of obesity-related metrics in cardiovascular risk assessments. These novel insights could lead to more personalized prevention strategies for individuals who may not exhibit classic cardiovascular risk profiles.

## Limitations

Despite the promising results, several limitations of the study need to be considered. First, the quality of the data, particularly with respect to missing values, poses a challenge. Although feature selection techniques, such as least absolute shrinkage and selection operator regression and SHAP analysis, were used to mitigate this, the impact of missing data remains a potential limitation. Second, the generalizability of the findings is limited because the study relies on a specific population. The results may not fully apply to populations with different demographic and clinical characteristics. To address this, future research should focus on evaluating these ML models in real-world clinical settings, where variability in clinical practice, missing data, and other factors may affect model performance.

## Conclusions

This study demonstrates the potential of ML in predicting CHD. The SHAP method enhances the interpretability of the prediction model, aiding health care professionals in clinical practice by supporting effective risk management and intervention strategies.

## Acknowledgments

This article was supported by the Japan Science and Technology Agency COI-NEXT (grant JPMJPF2018) to MA.

## Data Availability

The dataset examined in this study is not available to the public due to the inclusion of individuals' personal information but is available from the corresponding author at a reasonable request.

## Authors' Contributions

TV conceptualized and designed the study, conducted the data analysis and interpretation, and drafted the manuscript. Y Kokubo contributed to the study concept and design, curated the data, provided resources, and supervised the project. MA assisted in study design, data curation, and supervision. MI and MY contributed to data analysis and interpretation. RD, AMM, JTT, AA, MT, YMN, and TI provided critical feedback and revised the manuscript. All authors reviewed and approved the final version of the manuscript.

## Conflicts of Interest

YMN is employed by the Department of Digital Health and Epidemiology, Graduate School of Medicine and Public Health, Kyoto University, an Industry-Academia Collaboration Course supported by Eisai Co., Ltd. and Kyowa Kirin Co., Ltd. Additionally, YMN reports a study grant from Bayer outside the submitted work.

Multimedia Appendix 1
Percentage of missing data across all variables prior to imputation (Japanese participants, aged 30-84 years, Suita Study).
[PNG File, 93 KB - cardio_v9i1e68066_app1.png ]

Multimedia Appendix 2
Correlation Coefficients for variables used in CHD Incidence prediction (Japanese participants, aged 30-84 years, Suita Study).
[PNG File, 260 KB - cardio_v9i1e68066_app2.png ]

Multimedia Appendix 3
List of variables included in the CHD Incidence prediction model (Japanese participants, aged 30-84 years, Suita Study).
[DOCX File, 19 KB - cardio_v9i1e68066_app3.docx ]

Checklist 1
TRIPOD + AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis and Artificial Intelligence) checklist
[PDF File, 1663 KB - cardio_v9i1e68066_app4.pdf ]

## References


1. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9]
2. Roth GA, Mensah GA, Johnson CO, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019. J Am Coll Cardiol 2020 Dec;76(25):2982-3021. [doi: 10.1016/j.jacc.2020.11.010]
3. Lim HY, Burrell LM, Brook R, Nandurkar HH, Donnan G, Ho P. The need for individualized risk assessment in cardiovascular disease. J Pers Med 2022 Jul 14;12(7):1140. [doi: 10.3390/jpm12071140] [Medline: 35887637]
4. Matheson MB, Kato Y, Baba S, Cox C, Lima JAC, Ambale-Venkatesh B. Cardiovascular risk prediction using machine learning in a large Japanese cohort. Circ Rep 2022 Dec 9;4(12):595-603. [doi: 10.1253/circrep.CR-22-0101] [Medline: 36530840]
5. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE 2017;12(4):e0174944. [doi: 10.1371/journal.pone.0174944]
6. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. Behav Ther 2020 Sep;51(5):675-687. [doi: 10.1016/j.beth.2020.05.002] [Medline: 32800297]
7. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. Entropy (Basel) 2020 Dec 25;23(1):18. [doi: 10.3390/e23010018] [Medline: 33375658]
8. Vu T, Kokubo Y, Inoue M, et al. Machine learning approaches for stroke risk prediction: findings from the Suita Study. J Cardiovasc Dev Dis 2024 Jul 1;11(7):207. [doi: 10.3390/jcdd11070207] [Medline: 39057627]
9. Martin-Morales A, Yamamoto M, Inoue M, Vu T, Dawadi R, Araki M. Predicting cardiovascular disease mortality: leveraging machine learning for comprehensive assessment of health and nutrition variables. Nutrients 2023 Sep 11;15(18):3937. [doi: 10.3390/nu15183937] [Medline: 37764721]
10. Thanh NT, Luan VT, Viet DC, Tung TH, Thien V. A machine learning-based risk score for prediction of mechanical ventilation in children with dengue shock syndrome: a retrospective cohort study. PLoS ONE 2024;19(12):e0315281. [doi: 10.1371/journal.pone.0315281] [Medline: 39642139]
11. Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz Comput 2023;14(7):8459-8486. [doi: 10.1007/s12652-021-03612-z] [Medline: 35039756]
12. Forrest IS, Petrazzini BO, Duffy Á, et al. Machine learning-based marker for coronary artery disease: derivation and validation in two longitudinal cohorts. The Lancet 2023 Jan;401(10372):215-225. [doi: 10.1016/S0140-6736(22)02079-7]
13. Bani Hani SH, Ahmad MM. Machine-learning algorithms for ischemic heart disease prediction: a systematic review. Curr Cardiol Rev 2023;19(1):e090622205797. [doi: 10.2174/1573403X18666220609123053] [Medline: 35692135]
14. Alizadehsani R, Abdar M, Roshanzamir M, et al. Machine learning-based coronary artery disease diagnosis: a comprehensive review. Comput Biol Med 2019 Aug;111:103346. [doi: 10.1016/j.compbiomed.2019.103346] [Medline: 31288140]
15. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
16. Tunstall-Pedoe H, Kuulasmaa K, Amouyel P, Arveiler D, Rajakangas AM, Pajak A. Myocardial infarction and coronary deaths in the World Health Organization MONICA Project. Registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. Circulation 1994 Jul;90(1):583-612. [doi: 10.1161/01.cir.90.1.583] [Medline: 8026046]
17. Kokubo Y, Kamide K, Okamura T, et al. Impact of high-normal blood pressure on the risk of cardiovascular disease in a Japanese urban cohort: the Suita Study. Hypertension 2008 Oct;52(4):652-659. [doi: 10.1161/HYPERTENSIONAHA.108.118273] [Medline: 18725580]
18. Kokubo Y, Watanabe M, Higashiyama A, et al. Interaction of blood pressure and body mass index with risk of incident atrial fibrillation in a Japanese urban cohort: the Suita Study. Am J Hypertens 2015 Nov;28(11):1355-1361. [doi: 10.1093/ajh/hpv038] [Medline: 25845964]

19. Imai E, Horio M, Nitta K, et al. Estimation of glomerular filtration rate by the MDRD study equation modified for Japanese patients with chronic kidney disease. Clin Exp Nephrol 2007 Mar;11(1):41-50. [doi: 10.1007/s10157-006-0453-4] [Medline: 17384997]
20. Kokubo Y, Watanabe M, Higashiyama A, Nakao YM, Nakamura F, Miyamoto Y. Impact of intima-media thickness progression in the common carotid arteries on the risk of incident cardiovascular disease in the Suita Study. J Am Heart Assoc 2018 Jun 1;7(11):e007720. [doi: 10.1161/JAHA.117.007720] [Medline: 29858361]
21. Christodoulou E, van Smeden M, Edlinger M, et al. Adaptive sample size determination for the development of clinical prediction models. Diagn Progn Res 2021 Mar 22;5(1):6. [doi: 10.1186/s41512-021-00096-5] [Medline: 33745449]
22. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]
23. Bewick V, Cheek L, Ball J. Statistics review 14: logistic regression. Crit Care 2005 Feb;9(1):112-118. [doi: 10.1186/cc3045] [Medline: 15693993]
24. Su X, Xu Y, Tan Z, et al. Prediction for cardiovascular diseases based on laboratory data: an analysis of random forest model. J Clin Lab Anal 2020 Sep;34(9):e23421. [doi: 10.1002/jcla.23421] [Medline: 32725839]
25. Unnikrishnan P, Kumar DK, Poosapadi Arjunan S, Kumar H, Mitchell P, Kawasaki R. Development of health parameter model for risk prediction of CVD using SVM. Comput Math Methods Med 2016;2016:3016245. [doi: 10.1155/2016/3016245] [Medline: 27594895]
26. Son YJ, Kim HG, Kim EH, Choi S, Lee SK. Application of support vector machine for prediction of medication adherence in heart failure patients. Healthc Inform Res 2010 Dec;16(4):253-259. [doi: 10.4258/hir.2010.16.4.253] [Medline: 21818444]
27. Vu T, Dawadi R, Yamamoto M, et al. Prediction of depressive disorder using machine learning approaches: findings from the NHANES. BMC Med Inform Decis Mak 2025 Feb 17;25(1):83. [doi: 10.1186/s12911-025-02903-1] [Medline: 39962516]
28. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res 2019;3:18. [doi: 10.1186/s41512-019-0064-7] [Medline: 31592444]
29. Zhang Z, Rousson V, Lee WC, et al. Decision curve analysis: a technical note. Ann Transl Med 2018 Aug;6(15):308. [doi: 10.21037/atm.2018.07.02] [Medline: 30211196]
30. Bloch L, Friedrich CM, Alzheimer's Disease Neuroimaging Initiative. Data analysis with Shapley values for automatic subject selection in Alzheimer's disease data sets using interpretable machine learning. Alzheimers Res Ther 2021 Sep 15;13(1):155. [doi: 10.1186/s13195-021-00879-4] [Medline: 34526114]
31. Kawai T, Ohishi M, Takeya Y, et al. Carotid plaque score and intima media thickness as predictors of stroke and mortality in hypertensive patients. Hypertens Res 2013 Oct;36(10):902-909. [doi: 10.1038/hr.2013.61] [Medline: 23823172]
32. Nambi V, Chambless L, Folsom AR, et al. Carotid intima-media thickness and presence or absence of plaque improves prediction of coronary heart disease risk: the ARIC (Atherosclerosis Risk In Communities) study. J Am Coll Cardiol 2010 Apr 13;55(15):1600-1607. [doi: 10.1016/j.jacc.2009.11.075] [Medline: 20378078]
33. Ji C, Wang N, Shi J, et al. Level of systolic blood pressure within the normal range and risk of cardiovascular events in the absence of risk factors in Chinese. J Hum Hypertens 2022 Oct;36(10):933-939. [doi: 10.1038/s41371-021-00598-1] [Medline: 34480099]
34. Whelton SP, McEvoy JW, Shaw L, et al. Association of normal systolic blood pressure level with cardiovascular disease in the absence of risk factors. JAMA Cardiol 2020 Sep 1;5(9):1011-1018. [doi: 10.1001/jamacardio.2020.1731] [Medline: 32936272]
35. Li J, Somers VK, Gao X, et al. Evaluation of optimal diastolic blood pressure range among adults with treated systolic blood pressure less than 130 mm Hg. JAMA Netw Open 2021 Feb 1;4(2):e2037554. [doi: 10.1001/jamanetworkopen.2020.37554] [Medline: 33595663]
36. Guo LL, Chen YQ, Lin QZ, et al. Non-HDL-C Is more stable than LDL-C in assessing the percent attainment of non-fasting lipid for coronary heart disease patients. Front Cardiovasc Med 2021;8:649181. [doi: 10.3389/fcvm.2021.649181] [Medline: 33869310]
37. Saito I, Yamagishi K, Kokubo Y, et al. Non-high-density lipoprotein cholesterol and risk of stroke subtypes and coronary heart disease: The Japan Public Health Center-Based Prospective (JPHC) Study. JAT 2020;27(4):363-374. [doi: 10.5551/jat.50385]
38. Dong J, Yang S, Zhuang Q, et al. The associations of lipid profiles with cardiovascular diseases and death in a 10-year prospective cohort study. Front Cardiovasc Med 2021;8:745539. [doi: 10.3389/fcvm.2021.745539] [Medline: 34901209]
39. Zhao X, Wang D, Qin L. Lipid profile and prognosis in patients with coronary heart disease: a meta-analysis of prospective cohort studies. BMC Cardiovasc Disord 2021 Feb 3;21(1):69. [doi: 10.1186/s12872-020-01835-0] [Medline: 33535982]
40. Poznyak AV, Litvinova L, Poggio P, Sukhorukov VN, Orekhov AN. Effect of glucose levels on cardiovascular risk. Cells 2022 Sep 28;11(19):3034. [doi: 10.3390/cells11193034] [Medline: 36230996]
41. Riise HKR, Igland J, Sulo G, et al. Casual blood glucose and subsequent cardiovascular disease and all-cause mortality among 159 731 participants in Cohort of Norway (CONOR). BMJ Open Diabetes Res Care 2021 Feb;9(1):e001928. [doi: 10.1136/bmjdrc-2020-001928] [Medline: 33622686]
42. Selvin E, Coresh J, Golden SH, Brancati FL, Folsom AR, Steffes MW. Glycemic control and coronary heart disease risk in persons with and without diabetes. Arch Intern Med 2005 Sep 12;165(16):1910. [doi: 10.1001/archinte.165.16.1910]

43. Charoen P, Nitsch D, Engmann J, et al. Mendelian randomisation study of the influence of eGFR on coronary heart disease. Sci Rep 2016 Jun 24;6:28514. [doi: 10.1038/srep28514] [Medline: 27338949]
44. Jankowski J, Floege J, Fliser D, Böhm M, Marx N. Cardiovascular disease in chronic kidney disease. Circulation 2021 Mar 16;143(11):1157-1172. [doi: 10.1161/CIRCULATIONAHA.120.050686]
45. Brugts JJ, Knetsch AM, Mattace-Raso FUS, Hofman A, Witteman JCM. Renal function and risk of myocardial infarction in an elderly population: the Rotterdam Study. Arch Intern Med 2005;165(22):2659-2665. [doi: 10.1001/archinte.165.22.2659] [Medline: 16344425]
46. Maas A, Appelman YEA. Gender differences in coronary heart disease. Neth Heart J 2010 Dec;18(12):598-602. [doi: 10.1007/s12471-010-0841-y] [Medline: 21301622]
47. Shah T, Palaskas N, Ahmed A. An update on gender disparities in coronary heart disease care. Curr Atheroscler Rep 2016 May;18(5):28. [doi: 10.1007/s11883-016-0574-5] [Medline: 27029220]


## Abbreviations

**AUC:** area under the curve
**CCA:** common carotid artery
**CHD:** coronary heart disease
**DCA:** decision curve analysis
**eGFR:** estimated glomerular filtration rate
**HDL-c:** high-density lipoprotein cholesterol
**IMT:** intima-media thickness
**IMT_cMax:** maximum intima-media thickness of common carotid arteries
**LightGBM:** Light Gradient-Boosting Machine
**LR:** logistic regression
**MI:** myocardial infarction
**ML:** machine learning
**OE:** observed-to-expected
**RF:** random forest
**SBP:** systolic blood pressure
**SHAP:** Shapley Additive Explanation
**SVM:** support vector machine
**TRIPOD+AI:** Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis and Artificial Intelligence
**XGBoost:** Extreme Gradient Boosting


*Edited by A Coristine; submitted 29.10.24; peer-reviewed by M Gasmi, M Nomali, N Misra; revised version received 03.02.25; accepted 24.02.25; published 12.05.25.*

*Please cite as:*
*Vu T, Kokubo Y, Inoue M, Yamamoto M, Mohsen A, Martin-Morales A, Dawadi R, Inoue T, Tay JT, Yoshizaki M, Watanabe N, Kuriya Y, Matsumoto C, Arafa A, Nakao YM, Kato Y, Teramoto M, Araki M*
*Machine Learning Model for Predicting Coronary Heart Disease Risk: Development and Validation Using Insights From a Japanese Population–Based Study*
*JMIR Cardio 2025;9:e68066*
*URL: https://cardio.jmir.org/2025/1/e68066*
*doi:10.2196/68066*

# Pharmacist-Initiated Team-Based Intervention for Optimizing Guideline-Directed Lipid Therapy of Hospitalized Patients With Acute Coronary Syndrome: Pilot Study Using a Stepped-Wedge Cluster Design

Gayle L Flo[1], APRN, CNP; Mateo Alzate Aguirre[2,3], MD; Benjamin R Gochanour[4], MS; Kristin J Hynes[1,5], PharmD, RPh; Christopher G Scott[4], MS; Angela L Fink[6], APRN, CNP, MSN; Adelaide M Arruda-Olson[1], MD, PhD

[1]Department of Cardiovascular Medicine, Mayo Clinic, 200 First Street Southwest, Rochester, MN, United States

[2]Ascension St. John Hospital, Detroit, MI, United States

[3]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, United States

[4]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, United States

[5]Pharmacy Services, Mayo Clinic, Rochester, MN, United States

[6]Department of Nursing, Mayo Clinic, Rochester, MN, United States

**Corresponding Author:**
Adelaide M Arruda-Olson, MD, PhD
Department of Cardiovascular Medicine, Mayo Clinic, 200 First Street Southwest, Rochester, MN, United States

## Abstract

**Background:** Clinical guidelines recommend high-intensity statin therapy for patients with acute coronary syndrome (ACS). However, high-intensity statins have been underused in this population.

**Objective:** The objective of this study was to evaluate the feasibility of a pharmacist-initiated, team-based intervention for the delivery of individualized, guideline-directed, lipid-lowering therapy for patients with ACS.

**Methods:** Patients admitted with ACS to cardiology hospital services at Mayo Clinic from August 1, 2021, to June 19, 2022, were assigned to a pharmacist-initiated, team-based intervention group or control group using a stepped wedge cluster study design. For the intervention group, pharmacists reviewed electronic health records and provided recommendations for lipid lowering therapy in hospital and at follow-up. In the control group, patients received usual care. Neither care team, nor study team were blinded to study assignments. The primary outcome was the proportion of patients with ACS discharged on high-intensity statins in the intervention group compared to controls. Secondary outcomes were (1) proportion of patients in the intervention group with a specific templated pharmacist intervention note in their electronic health records, (2) frequency of low-density lipoprotein (LDL) measurements in hospital, (3) proportion of patients with information related to lipid follow-up in their discharge summary, and (4) proportion of patients that received LDL monitoring at the outpatient follow-up 4 to 12 weeks post discharge.

**Results:** There were 410 patients included in this study (median age 68, IQR 60-78 years) of whom 285 (69.5%) were male. Of the 402 patients alive at discharge, 355 (88.3%) were discharged taking a high-intensity statin, with no significant difference (*P*=.89) observed between groups. Lipid levels were measured in the hospital for 176/210 (83.8%) patients in the intervention group and 155/200 (77.5%) patients in the control group (*P*=.14). Fifty-four of 205 (26.3%) intervention patients alive at discharge had lipid-related recommendations in their discharge summary compared to 27/197 (13.7%) controls (*P*=.002). Forty-seven of 81 (58%) patients with lipid management recommendations provided in the discharge summary had LDL measured in the follow-up period compared with only 119/321 (37.1%) patients without these recommendations (*P*=.001). Of the 402 patients who survived to discharge, 166 (41.3%) had LDL measured at follow-up; the median LDL level was 63.5 (IQR 49-79) mg/dL, and distributions were similar by group (*P*=.95). Only 101/166 (60.8%) patients had follow-up LDL values below the target of 70 mg/dL.

**Conclusions:** During hospitalization, there was no group difference in the primary outcome of high-intensity statin therapy. Feasibility of an effective pharmacist-initiated intervention for improvement of lipid management was demonstrated by entry of recommendations in the discharge summary and related adjustment in outpatient statin therapy. The main opportunity for future improvement in lipid management of patients with ACS is in longitudinal patient follow-up.

***(JMIR Cardio 2025;9:e58837)*** doi:10.2196/58837

**KEYWORDS**

coronary disease; follow-up studies; lipids; myocardial infarction; statins

## Introduction

Acute coronary syndrome (ACS) includes non–ST-segment elevation myocardial infarction, ST-segment elevation myocardial infarction, and unstable angina [1-3]. Current estimates show approximately 605,000 new and 200,000 recurrent infarctions each year in the United States [4]. In 2020, there were 577,275 hospital discharges for ACS diagnosis [4]. Data from a Swedish registry revealed that approximately 20% of 97,254 patients who survived a myocardial infarction experienced another ischemic cardiac event within 24 months [5]. The 5-year mortality for ACS from large United Kingdom and Belgian studies ranged from 19% to 22% [6,7].

High-intensity statin therapy in the setting of ACS yields significant mortality benefit [8,9]. Hence, clinical practice guidelines recommend statin therapy for all patients with ACS [10,11]. In addition to decreasing low-density lipoprotein (LDL) levels, statins also promote improvement of endothelial function, decrease of platelet aggregation, and reduction of vascular inflammation [12]. LDL levels are used to monitor the intensity of therapy [13-15]. Guideline-directed therapies, including statins have been underused by patients with ACS [16]. For example, in a large cohort of 690,524 patients with recent ACS, less than half were on any statin therapy, and of those, only 20% were on high-intensity statins [17]. Another study which included 7802 patients with ACS showed that only one-third were prescribed a high-intensity statin at index hospitalization, and of those, only half were on such therapy at 1 year of follow-up [18].

Prior studies have demonstrated improved use of guideline-directed medical therapy by using team-based care delivery models. One prior study achieved sustained decreases in LDL levels to a specified target when pharmacists managed therapy for patients with coronary heart disease in the outpatient setting [19]. Another study showed that a pharmacist-initiated, team-based intervention with admission and predischarge medication reconciliation resulted in better adherence to guideline-directed therapy and reduced readmissions for heart failure [20]. The need to develop care delivery models to promote improved achievement of LDL targeted therapy is further supported by the work of Basaran et al [21] who analyzed data from 873 patients with diabetes from the EHPESUS registry which revealed that only 19.5% of the primary prevention and 7.5% of the secondary prevention groups were at LDL goal.

We hypothesize that a team-based inpatient care delivery model with processes that promote use of guideline-directed medical therapy for lipid management may improve outcomes for patients with ACS. An important unmet need exists to optimize lipid-lowering therapy for patients with ACS. Accordingly, the aim of this pilot study was to evaluate the feasibility of a pharmacist-initiated, team-based inpatient intervention for delivery of individualized, guideline-directed, lipid-lowering therapy recommendations for patients with ACS and to collect preliminary data on effectiveness.

## Methods

### Recruitment

This study was performed from August 1, 2021, to June 19, 2022, in 6 cardiology hospital services which admit patients with suspected ACS at Mayo Clinic in Rochester, Minnesota. Patients were included if they had a new diagnosis of ACS, that is, non–ST-segment elevation myocardial infarction, ST-segment elevation myocardial infarction, or unstable angina. Inclusion criteria remained consistent throughout the entire trial.

### Study Design

#### *Overview*

All patients admitted with ACS to cardiology were assigned to the control group (usual care) during the first 2 months of the project. At the beginning of month 3, the cardiology services began crossing over to the intervention group following a stepped wedge design [22] (Figure 1). Hence, each service had exposure to control status and intervention status over this study's period in longitudinal fashion. Each cluster of patients was unique in that patients with repeat admissions were excluded from this study at subsequent admissions. Neither the care team nor this study's team were blinded to the intervention status of patients.

**Figure 1.** Stepped wedge cluster allocation of patients.

| | Time, n (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Service** | Aug 1 to Oct 3 | Oct 4 to Nov 1 | Nov 2 to Dec 5 | Dec 6 to Jan 9 | Jan 10 to Feb 6 | Feb 7 to March 6 | March 7 to June 19 | **Total** |
| Cardiology 1 | 37 (9) | 13 (3.2) | 12 (2.9) | 13 (3.2) | 6 (1.5) | 3 (0.7) | 25 (6.1) | 109 (26.6) |
| Cardiology 2 | 14 (3.4) | 4 (1) | 8 (2) | 9 (2.2) | 13 (3.2) | 1 (0.2) | 13 (3.2) | 62 (15.1) |
| Cardiac Intensive Care Unit | 19 (4.6) | 29 (7.1) | 6 (1.5) | 9 (2.2) | 5 (1.2) | 7 (1.7) | 27 (6.6) | 102 (24.9) |
| Cardiology 4 | 15 (3.7) | 5 (1.2) | 13 (3.2) | 3 (0.7) | 6 (1.5) | 5 (1.2) | 21 (5.1) | 68 (16.6) |
| Cardiology 3 | 10 (2.4) | 5 (1.2) | 4 (1) | 5 (1.2) | 5 (1.2) | 1 (0.2) | 6 (1.5) | 36 (8.8) |
| Cardiology 5 | 6 (1.5) | 5 (1.2) | 6 (1.5) | 5 (1.2) | 4 (1) | 0 (0) | 7 (1.7) | 33 (8) |
| **Total** | 101 (24.6) | 61 (14.9) | 49 (12) | 44 (10.7) | 39 (9.5) | 17 (4.1) | 99 (24.1) | 410 (100) |

Baseline characteristics were collected for all patients enrolled. Data collection occurred via electronic health record (EHR) review after hospital admission with further completion of the datasets throughout this study's period. Statin therapy was defined as low-intensity (pravastatin, 10 and 20 mg; simvastatin, 10 mg), moderate-intensity (atorvastatin, 10 and 20 mg; pravastatin, 40 and 80 mg; rosuvastatin, 5 and 10 mg; and simvastatin 20 - 40 mg), or high-intensity (atorvastatin, 40 and 80 mg; rosuvastatin, 20 and 40 mg). Sample size calculations were not performed. The intent was to collect data for 8 months based on project timeline and resource allocation.

### *Control Group*

Patients in the control group received standard care for ACS, which included high-intensity statin therapy as recommended by clinical practice guidelines [10,11]. Each cardiology team was comprised of internal medicine residents and advanced practice providers (nurse practitioners or physician assistants) supervised by cardiologists. These teams collaborated with cardiology pharmacists who provided guidance about lipid therapy. All cardiology hospital pharmacists rotate covering each of the 6 services based on pre-established staffing schedules. The pharmacists were responsible for reviewing the patients' EHR daily, completing admission and discharge medication reconciliation, and entering recommendations. The pharmacists also rounded with hospital services to collaborate with the team regarding medication management.

### *Pharmacist-Initiated, Team-Based Intervention*

The primary objective of the pharmacist-initiated, team-based intervention was to ensure initiation or continuation of high-intensity statins, and the addition of ezetimibe if patients already taking a high-intensity statin had LDL level greater than 70 mg/dL on either most recent outpatient testing or in-hospital testing.

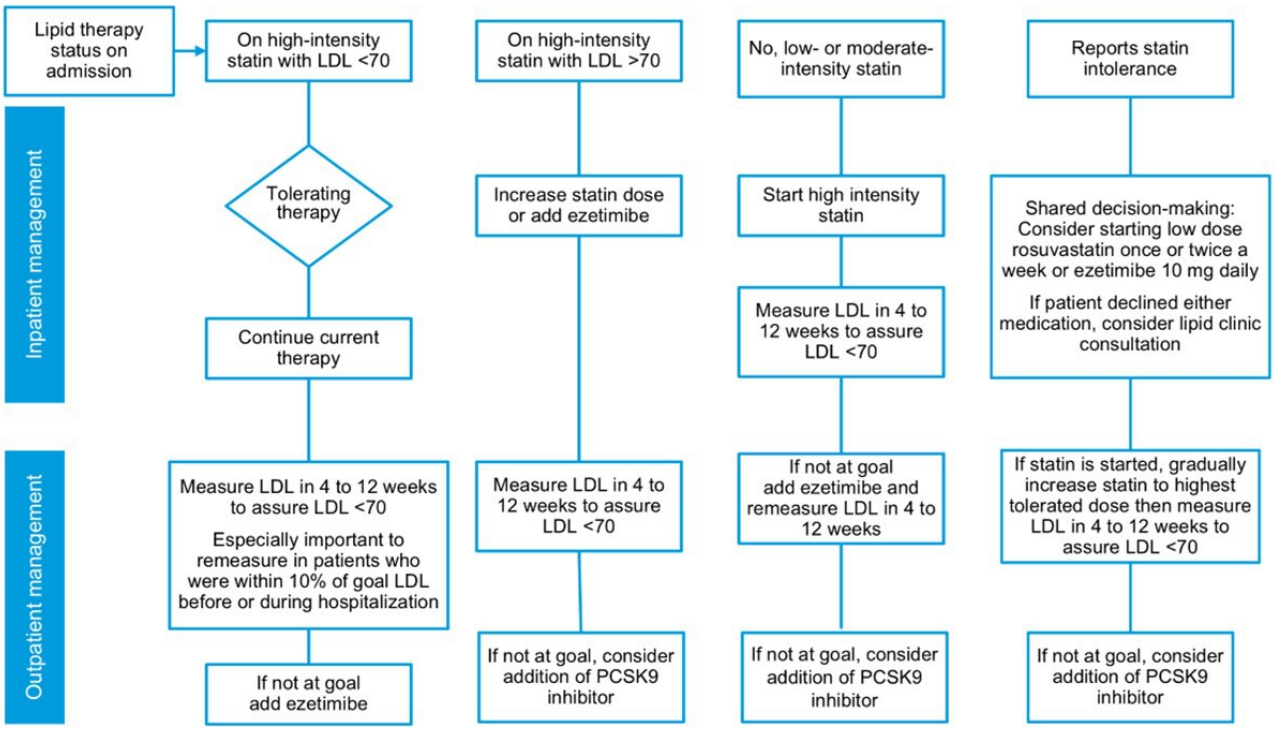
The cardiology pharmacist group consisted of 9 pharmacists who received training and instructions regarding implementation of the intervention in the form of presentations at staff meetings and written documents shared via emails describing project goals and pharmacist roles. At the beginning of each hospital service the cardiologists and team members entering the intervention phase received an email from this study's team describing the project.

After patients with ACS were admitted to the hospital, the pharmacists reviewed the EHR and interviewed each patient to gather information about adverse effects to statins and evaluate preadmission LDL levels from the EHR. Subsequently, contraindications to statins and adverse effects were documented in the pharmacist EHR note. If a lipid panel was not available from the prior 6 months, the pharmacists recommended checking a lipid panel to the cardiology team. After reviewing lipid levels, the pharmacists provided specific recommendations for the cardiology team members via EHR text messages and verbal communication.

The pharmacist recommendation algorithm is summarized in Figure 2. If the patient had an LDL<70 mg/dL and was on high-intensity statin, this medication was continued without change; if the LDL was >70 mg/dL while on a high-intensity statin the options were to increase statin dose or add ezetimibe. If the patient was not on a statin or was taking a moderate-intensity statin therapy, the moderate-intensity statin was discontinued and replaced by a high-intensity statin irrespective of LDL level. If the patient reported prior statin intolerance management options included (1) initiation of low-dose rosuvastatin 5 mg once or twice a week, (2) initiation

of ezetimibe only, or (3) patient referral for lipid clinic consultation at the lipid clinic. Each of these processes involved patient-centered shared decision-making for the selection of management strategy.

**Figure 2.** Pharmacist recommendation algorithm. LDL: low-density lipoprotein; PCSK9: proprotein convertase subtilisin/kexin type 9.




The pharmacists documented results of their review and recommendations in specially formatted pharmacist intervention notes. These notes recommended lipid testing within 4 to 12 weeks after discharge and treatment modifications if LDL remained greater than 70 mg/dL. Pharmacists requested that cardiology team members include these recommendations in discharge summaries sent to the primary care provider via the EHR. Fidelity with the intervention was evaluated by the presence and content of a templated pharmacist intervention note documented in the EHR.

The pharmacist notes advised repeat lipid measurements at 4 to 12 weeks after hospital discharge, as recommended by the guidelines [10]. However, very few patients underwent testing within 12 weeks. Therefore, the data collection interval was extended to 6 months post hospital discharge. The low frequency of testing by 12 weeks was likely related to clinical decisions and appointment availability in the outpatient clinics. The research team had no influence on scheduling of follow-up appointments.

Follow-up outcomes were obtained by manual review of the EHR within 6 months of hospital discharge. Variables obtained at follow-up were LDL results, test date, and adjustments in lipid therapy made at follow-up. A REDCap (Research Electronic Data Capture, Vanderbilt University 2022) database and Microsoft Excel (Microsoft Corp) were used for data entry and storage.

### *Outcomes*

The primary outcome was the proportion of patients with ACS discharged on high-intensity statins in the intervention group compared to the control group. Secondary outcomes were (1) proportion of patients in the intervention group with the specific templated pharmacist intervention note in the EHR, (2) frequency of LDL measurements in the hospital, (3) proportion of patients with information related to lipid follow-up in their discharge summary, and (4) proportion of patients that received LDL monitoring at outpatient follow-up 4 to 12 weeks post discharge.

### *Statistical Methods*

Baseline demographic characteristics of the patients were summarized as median (IQR) for continuous and count (proportion expressed as percentage) for categorical variables. Baseline comparisons of continuous variables between groups were made with the Wilcoxon rank-sum test, and comparisons of categorical variables were made with the chi-square or Fisher exact tests.

Preadmission and in-hospital LDL levels were compared by a paired *t* test (2-tailed). The $\chi^2$ test was used to assess impact of the intervention on the number of patients who had lipid levels measured during hospitalization and the percentage of patients discharged on high-intensity statin therapy. The effect of the intervention on changes made in lipid-lowering therapy from admission to discharge was assessed using a Cochran-Mantel-Haenszel test. Overall rates of admission without lipid therapy compared to discharge without lipid therapy were evaluated by the McNemar test. LDL levels at follow-up were compared by group with the unpaired *t* test. Other follow-up outcome comparisons were made using the $\chi^2$ test.

A stepped wedge cluster design [22] was used for subject allocation, with the cardiology services as clusters (Figure 1). We evaluated the effects of admission period and cardiology service (rows and columns of Figure 1, respectively) on the outcome variables of interest and found that the results are not likely confounded by these factors. This evaluation was initially performed visually. Subsequently variables were added as covariates in the regression models. No significance or discernable patterns were found; therefore, only the simplified (unadjusted) results are presented herein. For continuous variables, 95% CIs were computed using the normal approximation, and the CIs for binomial proportions were computed by the Wilson score method [23].

Both intent-to-treat (intended participant assignment based on stepped wedge design) and subgroup (intervention received vs all controls) analyses were conducted for groupwise differences, including when comparing rates of lipid measurements in hospital and rates of discharge on high-intensity statins. Analyses evaluating discharge and follow-up outcomes excluded patients who died during hospitalization. A 2-sided *P* value of <.05 was considered statistically significant. All statistical analyses were conducted using R (version 4.1.2 software; R Core Team, R Foundation).

### Ethical Considerations

This quality improvement study was approved by the Mayo Clinic Institutional Review Board (file 21‐009289). All patients agreed to have their medical records used for research, and the institutional review board waived the need for informed consent. Subject data were deidentified in all analysis files and have been password protected within the institutional fire walls. No compensation was provided to study participants.

## Results

### Cohort Characteristics and Intervention Delivery

A total of 410 patients admitted with ACS were included in this study. Of these, 200 patients were assigned to the control group and 210 to the intervention group (Table 1). Most patients were men (285/410, 69.5%), and the overall median age at admission was 68 (IQR 60-78) years. Patients in the intervention group were slightly older than those in the control group. The most frequent ACS diagnosis was non–ST-segment elevation myocardial infarction. Unstable angina represented a greater proportion of ACS diagnoses in the intervention group than the control group. Statin use at admission was similar across this study's groups, and almost half of patients were not taking statin medications at hospital admission. The pharmacists determined that 21/410 (5.1%) patients were not taking statin therapy due to prior intolerance, 120/410 (29.3%) patients were not taking statins because therapy had not been recommended, and 27/410 (6.6%) patients had previously declined statin therapy.

**Table .** Clinical characteristics of the cohort.

| Characteristic | Control group (n=200) | Intervention group (n=210) | *P* value |
|---|---|---|---|
| Age (years), median (IQR) | 66.5 (59 - 77) | 71 (61 - 79.8) | .02[a] |
| Sex, n (%) | | | |
| Male | 137 (68.5) | 148 (70.5) | .66[b] |
| Female | 63 (31.5) | 62 (29.5) | |
| Admitting ACS[c] diagnosis, n (%) | | | .003[b] |
| STEMI[d] | 56 (28) | 58 (27.6) | |
| NSTEMI[e] | 141 (70.5) | 137 (65.2) | |
| Unstable angina | 1 (0.5) | 15 (7.1) | |
| Other (troponin elevation) | 2 (1) | 0 (0) | |
| Admission therapy, n (%) | | | .51[b] |
| High-intensity statin[f] | 56 (28) | 65 (31) | |
| Moderate-intensity statin | 52 (26) | 46 (21.9) | |
| Low-intensity statin | 8 (4) | 5 (2.4) | |
| Nonstatin therapies | 3 (1.5) | 7 (3.3) | |
| No lipid-lowering therapy | 81 (40.5) | 87 (41.4) | |
| Inpatient LDL[g] level (mg/dL), median (IQR) | 93 (60 - 127.5) | 93.5 (63 - 130) | .70[a] |
| Missing, n | 45 | 34 | |
| Preadmission triglyceride level (mg/dL), median (IQR; within 6 mo) | 126 (90 - 183.8) | 149 (105.5 - 215.5) | .02[a] |
| Missing, n | 74 | 59 | |
| Prior diagnosis of hyperlipidemia, n (%) | 173 (88.3) | 180 (85.7) | .45[b] |
| Missing, n | 4 | 0 | |
| Prior diagnosis of hypertriglyceridemia, n (%) | 72 (42.9) | 104 (58.4) | .004[b] |
| Missing, n | 32 | 32 | |
| Prior diagnosis of diabetes, n (%) | 83 (41.7) | 79 (38) | .44[b] |
| Missing, n | 1 | 2 | |
| Prior diagnosis of hypertension, n (%) | 145 (72.5) | 147 (70.7) | .68[b] |
| Missing, n | 0 | 2 | |
| Length of hospital stay (days), median (IQR) | 3.5 (2 - 10) | 4 (2 - 9) | .96[a] |
| In-hospital deaths, n (%) | 3 (1.7) | 5 (2.4) | .73[h] |
| Missing, n | 19 | 3 | |
| Left ventricular ejection fraction, median (IQR) | 52 (38.8 - 60) | 55 (44 - 61) | .04[a] |
| Missing, n | 4 | 4 | |
| Comorbidities, n (%) | | | |
| Prior myocardial infarction | 34 (17.7) | 27 (13.3) | .23[b] |
| Missing, n | 8 | 7 | |
| Prior CABG[i] | 14 (7.1) | 23 (11.2) | .15[b] |
| Missing, n | 3 | 5 | |
| Prior PCI[j] | 63 (31.7) | 54 (26.1) | .22[b] |

| Characteristic | Control group (n=200) | Intervention group (n=210) | *P* value |
|---|---|---|---|
| Missing, n | 1 | 3 | |
| Prior diagnosis of heart failure | 43 (21.5) | 35 (16.7) | .22[b] |
| Missing, n | 0 | 1 | |
| Prior diagnosis of peripheral artery disease | 17 (8.6) | 26 (12.4) | .21[b] |
| Missing, n | 2 | 1 | |
| Prior ischemic stroke | 11 (5.6) | 14 (6.7) | .64[b] |
| Missing, n | 2 | 0 | |

[a]Wilcoxon rank-sum test.

[b]Pearson chi-square test.

[c]ACS: acute coronary syndrome.

[d]STEMI: ST-segment elevation myocardial infarction.

[e]NSTEMI: non–ST-segment elevation myocardial infarction.

[f]See methods section for definitions of statin intensity.

[g]LDL: low-density lipoprotein.

[h]Fisher exact test.

[i]CABG: coronary artery bypass grafting.

[j]PCI: percutaneous coronary intervention.

Preadmission LDL test results were available for 272/410 (66.3%) participants. The median preadmission LDL was 93 (IQR 63-134) mg/dL and did not differ significantly between groups. The distribution of hyperlipidemia, hypertension, and diabetes was also similar. However, patients in the intervention group were more likely to have prior diagnosis of elevated triglycerides and slightly higher levels of preadmission triglycerides.

The median length of hospitalization was 4 (IQR 2-9) days, which was similar across this study's groups. During hospitalization, 8 patients died, and the distribution of deaths was similar across study groups. Deaths were attributed to complications of acute myocardial infarction, including cardiogenic shock, respiratory failure from volume overload, or multisystem organ failure from persistent hypotension. The distribution was similar across this study's groups for left ventricular ejection fraction, prior myocardial infarction, history of coronary artery bypass grafting, percutaneous coronary intervention, peripheral arterial disease, and ischemic stroke.

To assign recommendations, the pharmacists categorized patients into the following groups: taking a high-intensity statin, had a recent LDL less than 70 mg/dL; taking a high-intensity statin, had a recent LDL more than 70 mg/dL; taking a high-intensity statin, no evidence of a recent LDL measurement; taking low- to moderate-intensity statin therapy; taking lipid-lowering therapy other than a statin; and not taking lipid lowering therapy. Table 2 shows prehospital statin dosing cross-referenced with LDL values. The proportion of patients in these subgroups was not significantly different (*P*=.49).

Among the 402 patients alive at hospital discharge, the proportion of patients taking a high-intensity statin increased significantly (*P*<.001) compared with admission proportions (121/402, 30.1% to 355/402, 88.3%) including 182/205 (88.8%, 95% CI 83.4% - 92.6%) intervention participants (intent-to-treat group) and 173/197 (87.8%, 95% CI 82.2% - 91.9%) control participants (*P*=.89; Table 3). When the subgroup that received the intervention (n=100) was compared to all controls, the findings were similar.

**Table .** Prehospital statin therapy and low-density lipoprotein (LDL) levels of patients taking lipid-lowering therapy.[a]

| Admission therapy and prehospital LDL level | Control group (n=200), n (%) | Intervention group (n=210), n (%) |
|---|---|---|
| HIS[b] with LDL≤70 mg/dL | 26 (13) | 23 (11) |
| HIS with LDL>70 mg/dL | 20 (10) | 29 (13.8) |
| HIS with no recent LDL measurement | 10 (5) | 13 (6.2) |
| Low- to moderate-intensity statin | 60 (30) | 51 (24.3) |
| Nonstatin therapy | 3 (1.5) | 7 (3.3) |
| No lipid therapy | 81 (40.5) | 87 (41.4) |

[a]The difference between groups was not statistically significant (*P*=.49).

[b]HIS: high-intensity statin.

**Table .** Admission and discharge medications among nondeceased patients.

| Treatment | Control group (n=197), n (%) | Intervention group (n=205), n (%) |
|---|---|---|
| Admission therapy | | |
| No lipid therapy | 80 (40.6) | 85 (41.5) |
| Nonstatin | 3 (1.5) | 6 (2.9) |
| Low-intensity statin | 8 (4.1) | 5 (2.4) |
| Moderate-intensity statin | 50 (25.4) | 44 (21.5) |
| High-intensity statin | 56 (28.4) | 65 (31.7) |
| Discharge therapy | | |
| No lipid therapy | 4 (2) | 4 (2) |
| Nonstatin | 4 (2) | 4 (2) |
| Low-intensity statin | 0 (0) | 3 (1.5) |
| Moderate-intensity statin | 16 (8.1) | 12 (5.9) |
| High-intensity statin | 173 (87.8) | 182 (88.8) |

Importantly, among patients admitted who were not receiving lipid lowering therapy, most (146/165, 88.5%) were taking a statin at discharge, and almost all patients taking a high-intensity statin at admission were taking a high-intensity statin at discharge (120/121, 99.2%). Eight patients were discharged without lipid therapy for the following reasons: 1 patient reported statin intolerance and recommendations were made to consider outpatient PCSK9 (proprotein convertase subtilisin/kexin type 9) inhibitor therapy; 1 patient had a non-ACS diagnosis at discharge, and statin therapy was appropriately withheld; 1 patient had an extremely low LDL level and preferred not to take a statin at hospital discharge; and 5 patients were discharged to hospice care and given comfort care.

The intervention was implemented for only 100/210 (47.6%) patients allocated to the intervention group, as indicated by inclusion of the templated pharmacist intervention note. Of these patients, 2 died in the hospital and 8 had recommendations coded as “other.” The pharmacist recommendations were followed (measured by the discharge medication) for 85 of the remaining 90 patients (94.4%, 95% CI 86.9% - 97.9%). See Table 4 for additional details.

**Table .** Pharmacist recommendations and inpatient low-density lipoprotein (LDL) measurement.

| Type of delivery recommendation | Control group (n=200), n (%) | Intervention group (n=210), n (%) | *P* value[a] |
|---|---|---|---|
| Type of pharmacist EHR[b] note | | | <.001 |
| Intervention and routine notes | 0 (0) | 9 (4.3) | |
| Intervention note only | 3 (1.5) | 91 (43.3) | |
| No note or note without lipid therapy recommendation | 114 (57) | 62 (29.5) | |
| Routine notes only | 83 (41.5) | 48 (22.9) | |
| Intervention assigned and received | | | |
| Yes | N/A[c] | 100 (47.6) | |
| Pharmacist recommendation | | | <.001 |
| Continue current statin | 16 (8) | 19 (9) | |
| Continue high-intensity statin, add ezetimibe | 2 (1) | 5 (2.4) | |
| Change from admission high-intensity statin to alternative high-intensity statin | 2 (1) | 2 (1) | |
| Recommend increase in high-intensity statin dose | 6 (3) | 12 (5.7) | |
| Begin low- to moderate-intensity statin | 2 (1) | 1 (0.5) | |
| Begin high-intensity statin | 25 (12.5) | 61 (29) | |
| Begin high-intensity statin and ezetimibe | 3 (1.5) | 0 (0) | |
| Change from low- to moderate-intensity statin to a high-intensity statin | 24 (12) | 30 (14.3) | |
| No note or note without recommendation | 114 (57) | 62 (29.5) | |
| Other[d] | 6 (3) | 18 (8.6) | |
| Inpatient LDL measured | 155 (77.5) | 176 (83.8) | .14 |

[a]Pearson chi-square test.

[b]EHR: electronic health record.

[c]N/A: not applicable.

[d]Other recommendations included alternative dosing and or drug due to past statin intolerance (12 patients), recommendation to start nonstatin therapy (2 patients), transition to hospice care (1 patient), remainder were variations due to coding interpretations (9 patients).

The intent-to-treat analysis showed that 176/210 (83.8%, 95% CI 78% - 88.4%) patients in the intervention group had lipid levels measured in the hospital compared with 155/200 (77.5%, 95% CI 71% - 83%) patients in the control group (*P*=.14; Table 4). The subgroup analysis yielded a similar, nonsignificant finding (87/100, 87% vs 155/200, 77.5%; *P*=.07). Among patients who had both before and after admission LDL levels measured, their mean in-hospital LDL levels were approximately 13 mg/dL lower than they were before hospitalization (95% CI −17.9 to −7.5; *P*<.001).

## Follow-Up Period Results

Patients randomized to the intervention group were more likely to have lipid management recommendations added to the discharge summary (54/205, 26.3% vs 27/197, 13.7%; *P*=.002). Subgroup analysis showed a stronger effect, with 38/98 (38.8%) patients who received the intervention having a lipid management recommendation in their discharge summary versus 27/197 (13.7%) controls (*P*<.001). More than half (47/81, 58%) of patients with the lipid management recommendations provided in the discharge summary had LDL measured in the follow-up period compared with only 119/321 (37.1%) patients without these recommendations (*P*=.001).

Documented LDL levels within 4 weeks to 6 months of hospital discharge were available for 166/402 (41.3%) patients and included 90/205 (43.9%) of intervention patients and 76/197 (38.6%) control patients (*P*=.33; Table 5). Among the 166 patients with LDL measurements, 101 (60.8%) had a follow-up LDL of less than 70 mg/dL (median 63.5, IQR 49-79 mg/dL). The median LDL for the control group was 63 (IQR 49-79)

mg/dL and for the intervention group 63.5 (IQR 49-78) mg/dL (*P*=.95). The subgroup analysis resulted in comparable findings.

**Table .** Low-density lipoprotein (LDL) assessment after patient discharge.[a]

| LDL | Control group (n=197) | Intervention group (n=205) | *P* value |
|---|---|---|---|
| LDL measured within 4 weeks to 6 months after discharge, n (%) | 76 (38.6) | 90 (43.9) | .33[b] |
| LDL values (mg/dL), median (IQR) | 63 (49 - 79) | 63.5 (49-78) | .95[c] |

[a]The 8 patients who died were excluded.

[b]Pearson chi-square test.

[c]Wilcoxon rank-sum test.

## Discussion

### Principal Findings

In the intervention group of this pilot study, pharmacists provided patient-centered recommendations for guideline-directed statin therapy for patients with ACS. At hospital discharge patients in both the intervention and controls groups had very high rates of statin therapy, such that there was no significant difference for the primary outcome. However, there was significant differences in the rates of pharmacist recommendations being incorporated into the discharge summary for the intervention group and these recommendations were associated with higher rates of adjustment of statin therapy at outpatient patient follow-up. These findings demonstrate feasibility for implementation and effectiveness of the in-hospital pharmacist intervention.

The rates for patients taking a high-intensity statin were high in both the intervention and control groups. The change in therapy from admission to discharge was significant; all patients eligible and consenting to statin therapy were discharged with high-intensity therapy.

A stepped wedge cluster study design was used due to logistical constraints [22] as subjects were recruited from 6 different cardiology hospital services. These services served as natural clusters for which we delivered the intervention. Additionally, by implementing the intervention within these clusters, both the staff training and deployment of the intervention were possible. Intervention fidelity was determined by the presence of the templated pharmacist intervention note in the EHR. We found that only 100/210 (47.6%) intervention patients had this type of note documented. During this pilot, the pharmacists were not assigned to a particular service but rather served patients across multiple services. This meant pharmacists sometimes cared for both control and intervention patients in the same day, increasing the risk of low intervention fidelity (intervention patients not receiving) or intervention contamination (controls receiving the intervention). While intervention fidelity was low, there were only 3 instances of intervention templated pharmacist notes appearing in the record for a control patient demonstrating low rate of contamination.

The estimated rate of in-hospital LDL measurement was similar between this study's groups. In both groups adherence to measuring LDL levels during hospitalization was high minimizing the opportunity to show improvement as a result of the intervention. LDL levels during hospitalization for ACS were lower than levels that were obtained within 6 months before the hospitalization for ACS event. Despite many patients having an in-patient LDL of 70 mg/dL or less during hospitalization, levels should be checked at follow-up post hospitalization as dose adjustments may be necessary. Overall, there was no difference in post hospitalization lipid measurement between the control group and the intervention group. However, intervention patients were more likely to have lipid therapy follow-up recommendations in their discharge summary, although rates were low in both groups. The subset of patients that had pharmacist recommendations for lipid testing available in the discharge summary had higher frequency of post hospital lipid measurement (*P*=.001). This suggests that communication of pharmacists' recommendations for outpatient providers delivered via discharge summaries was beneficial, indicating that pharmacists may have an important role in bridging the gap in guideline directed care between in-hospital and outpatient care [19].

The intervention proposed herein focused on recommendations for guideline-directed optimal lipid lowering medical therapy. Diet and lifestyle modifications are also important in lipid optimization and these recommendations are routinely provided for each patient during the hospitalization by the multidisciplinary care teams. Additionally, at hospital discharge patients with ACS are routinely referred to cardiac rehabilitation programs which include comprehensive cardiovascular health assessment as well as detailed recommendations for diet and physical activity [11].

### Comparison to Previous Work

Prior studies have demonstrated a strong correlation between statin intensity and survival of patients with ACS [9]. High-intensity statins have a significant impact on survival over moderate-intensity statins regardless of patient age [9]. For this reason, our clinical practice standard is to initiate high-intensity statins on all patients hospitalized with ACS. Low use of high-intensity statins post-ACS and difficulty achieving goal LDL levels may have a negative impact on secondary prevention in patients with ACS [9].

In a prior study it was demonstrated that high-intensity statin use increased from 33.5% to 71.7% among 117,989 patients discharged from the hospital after a myocardial infarction [24]. In that same study, older age, previous statin intolerance, drug

interactions, and long-term care goals were reasons that statins were not prescribed at discharge. This study showed high frequency of high-intensity statin prescription at hospital discharge, with the main reason that patients did not take statins being discharge to hospice for end-of-life care.

Previous studies demonstrated that in-hospital and follow-up lipid testing was associated with higher rates of lipid lowering therapy prescription for patients with ACS [25,26]. In this study herein, a lipid therapy recommendation in the discharge summary was associated with higher frequency of lipid testing during the follow-up period. In this study only 41% of all study patients had LDL measurements within 6 months of hospital discharge. Of these patients, 61% had an LDL less than 70 mg/dL hence nearly 40% of these patients with ACS who had follow-up lipid testing were not at goal LDL. This low frequency of follow-up lipid testing is not unique to our practice. Wang et al [27] compared data from 11,046 patients aged older than 65 years discharged from the hospital being alive from the years 2007 to 2009. In this cohort, only 44% had repeat lipid testing at 90 days and only 14% were on high-intensity statins at 1 year follow-up.

These studies highlight the need to implement interventions that improve use of lipid follow-up testing for the achievement of target LDL levels. Our proposed intervention promotes improved communication among providers including pharmacist recommendations shared across the continuum of care targeting lipid lowering therapy.

### Strengths and Limitations

The primary strength of this study is the ability to demonstrate alignment with guideline-directed high-intensity statin therapy for patients with ACS, while no overall group differences were seen this study identified an important opportunity for improved longitudinal lipid lowering therapy after hospital discharge in this high-risk population. This study suggests that a team-based approach may be successful and warrants further investigation and refinement.

This study has limitations. First, this pilot study was not randomized due to limited availability of clinical resources during this study's period. Randomization will be used in a larger implementation trial which will be endorsed by administrative leadership for coordination and allocation of clinical resources. Second, the intervention fidelity was low, potentially diluting the treatment effect and reducing sample size for the subgroup analysis of patients who received the intervention. This reduced sample size limited statistical power for detecting group differences. There are several potential causes for the observed low intervention fidelity. A new hospital wide pharmacy initiative for documentation of pharmacist progress notes in the EHR on all patients started during this pilot. Additionally, some patients were discharged from the hospital within 24 hours after admission, which decreased the opportunity for the pharmacists to deliver the intervention. In the future, we plan to schedule activation of the intervention for a time that does not overlap with other institutional quality initiatives and improve integration of the intervention with discharge planning. Lastly, the same pharmacists were responsible for covering multiple services and sometimes cared for intervention and control patients on the same day. In the future, we plan to clearly label in the EHR which group a given patient is assigned (control vs intervention) and when possible, assign different pharmacists for control versus intervention groups. By improving intervention fidelity, statistical power for detecting group differences may also improve.

Results of this study may be generalized to other clinical settings which use team-based care in hospital practice. The institution in which this project was performed is a referral institution which may have impacted the patient population characteristics, but the care delivered was guideline-based which should be adopted in all institutions caring for patients with ACS.

### Future Directions

Shortly after this pilot study was completed, an Expert Consensus paper was published by the American College of Cardiology recommending a target LDL for high-risk (including post-ACS) patients of less than 55 mg/dL [28]. The primary driver behind this consensus document was the availability of nonstatin therapies that can further help optimize LDL levels [6]. With lower target LDL levels and the advent of nonstatin lipid lowering therapies, the proposed intervention could be adapted to lower target LDL levels and the use of both statins and nonstatin lipid lowering therapies to promote the delivery of guideline-directed care for patients with ACS.

Multidisciplinary care processes that enhance best practices for lipid management after hospital discharge of patients with ACS are needed to improve patient outcomes. A previously published study from our institution described a proactive model of care delivery assisted by clinical decision support technology to promote delivery of guideline-directed care after patients are discharged from the hospital [9]. We envision implementation of a combined process of using the pharmacist-initiated program for lipid lowering therapy in the hospital setting and a proactive outpatient model of care delivery supported by technology as described by Partogi et al [29] to promote longitudinal patient follow-up for delivery of secondary prevention guideline-directed therapy for patients with ACS.

### Conclusions

An inpatient pharmacist-initiated intervention for lipid lowering therapy for patients with ACS is feasible and effective. The main opportunity for future improvement lies in improved communication via the EHR to promote optimization of lipid management in longitudinal outpatient follow-up in this population.

## Acknowledgments

Marianne Mallia, senior scientific and medical editor, Mayo Clinic, and Lyle J Olson, MD, substantively edited this paper. The Scientific Publications staff at Mayo Clinic provided proofreading, administrative, and clerical support. The authors thank Kara

M Firzlaff, administrative assistant, for secretarial support. This project was funded by a grant from the Mayo Clinic Department of Cardiovascular Medicine and the Mayo Clinic Center for Clinical and Translational Science (grant UL1TR002377). The content is solely responsibility of authors and does not necessarily represent official views of National Institutes of Health. No generative artificial intelligence resources were used in the creation of this paper.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

GF handled the conceptualization, investigation, project administration, writing of original draft, and review and editing. MAA worked on the conceptualization, data curation, investigation, visualization, and writing of the original draft. BG assisted with the data curation, and did the formal analysis, visualization, writing of the original draft, and review and editing. KH assisted with the conceptualization, investigation, project administration, writing of the original draft, and review and editing. CS aided with the formal analysis and supervision. AF helped with the conceptualization and funding acquisition. AMA-O carried out the conceptualization, methodology, supervision, writing of the original draft, and review and editing.

## Conflicts of Interest

None declared.

## References

1. Sweis RN, Jivan A. Overview of acute coronary syndromes. MSD Manual Professional Version. 2024. URL: https://www.msdmanuals.com/professional/cardiovascular-disorders/coronary-artery-disease/unstable-angina [accessed 2025-03-04]
2. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). Circulation 2018 Nov 13;138(20):e618-e651. [doi: 10.1161/CIR.0000000000000617] [Medline: 30571511]
3. Jaffe AS, Jaffe HA. Use of high-sensitivity cardiac troponin in patients with chronic comorbidities: moving from theory to practice. J Am Coll Cardiol 2018 Sep 4;72(10):1138-1140. [doi: 10.1016/j.jacc.2018.06.052] [Medline: 30165985]
4. Martin SS, Aday AW, Almarzooq ZI, et al. 2024 heart disease and stroke statistics: a report of US and global data from the american heart association. Circulation 2024 Feb 20;149(8):e347-e913. [doi: 10.1161/CIR.0000000000001209] [Medline: 38264914]
5. Jernberg T, Hasvold P, Henriksson M, Hjelm H, Thuresson M, Janzon M. Cardiovascular risk in post-myocardial infarction patients: nationwide real world data demonstrate the importance of a long-term perspective. Eur Heart J 2015 May 14;36(19):1163-1170. [doi: 10.1093/eurheartj/ehu505] [Medline: 25586123]
6. Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. N Engl J Med 2017 May 4;376(18):1713-1722. [doi: 10.1056/NEJMoa1615664] [Medline: 28304224]
7. Giberson SF. Million Hearts(TM): pharmacist-delivered care to improve cardiovascular health. Public Health Rep 2013;128(1):2-6. [doi: 10.1177/003335491312800102] [Medline: 23277654]
8. Murphy SA, Cannon CP, Wiviott SD, et al. Effect of intensive lipid-lowering therapy on mortality after acute coronary syndrome (a patient-level analysis of the Aggrastat to Zocor and Pravastatin or Atorvastatin evaluation and infection therapy-thrombolysis in myocardial infarction 22 trials). Am J Cardiol 2007 Oct 1;100(7):1047-1051. [doi: 10.1016/j.amjcard.2007.04.053] [Medline: 17884359]
9. Rodriguez F, Maron DJ, Knowles JW, Virani SS, Lin S, Heidenreich PA. Association between intensity of statin therapy and mortality in patients with atherosclerotic cardiovascular disease. JAMA Cardiol 2017 Jan 1;2(1):47-54. [doi: 10.1001/jamacardio.2016.4052] [Medline: 27829091]
10. Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary: a report of the american college of cardiology/american heart association task force on clinical practice guidelines. Circulation 2019 Jun 18;139(25):e1046-e1081. [doi: 10.1161/CIR.0000000000000624] [Medline: 30565953]
11. Smith SC Jr, Benjamin EJ, Bonow RO, et al. AHA/ACCF secondary prevention and risk reduction therapy for patients with coronary and other atherosclerotic vascular disease: 2011 update: a guideline from the American Heart Association and American College of Cardiology Foundation endorsed by the World Heart Federation and the Preventive Cardiovascular Nurses Association. J Am Coll Cardiol 2011 Nov 29;58(23):2432-2446. [doi: 10.1016/j.jacc.2011.10.824] [Medline: 22055990]
12. Oesterle A, Laufs U, Liao JK. Pleiotropic effects of statins on the cardiovascular system. Circ Res 2017 Jan 6;120(1):229-243. [doi: 10.1161/CIRCRESAHA.116.308537] [Medline: 28057795]

13. Yu S, Jin J, Chen Z, Luo X. High-intensity statin therapy yields better outcomes in acute coronary syndrome patients: a meta-analysis involving 26,497 patients. Lipids Health Dis 2020 Aug 23;19(1):194. [doi: 10.1186/s12944-020-01369-6] [Medline: 32829708]
14. Sposito AC, Chapman MJ. Statin therapy in acute coronary syndromes: mechanistic insight into clinical benefit. Arterioscler Thromb Vasc Biol 2002 Oct 1;22(10):1524-1534. [doi: 10.1161/01.atv.0000032033.39301.6a] [Medline: 12377727]
15. Hulten E, Jackson JL, Douglas K, George S, Villines TC. The effect of early, intensive statin therapy on acute coronary syndrome: a meta-analysis of randomized controlled trials. Arch Intern Med 2006 Sep 25;166(17):1814-1821. [doi: 10.1001/archinte.166.17.1814] [Medline: 17000936]
16. Krumholz HM, Normand SLT, Wang Y. Twenty-year trends in outcomes for older adults with acute myocardial infarction in the United States. JAMA Netw Open 2019 Mar 1;2(3):e191938. [doi: 10.1001/jamanetworkopen.2019.1938] [Medline: 30874787]
17. Klimchak AC, Patel MY, Iorga Ş, Kulkarni N, Wong ND. Lipid treatment and goal attainment characteristics among persons with atherosclerotic cardiovascular disease in the United States. Am J Prev Cardiol 2020 Mar;1:100010. [doi: 10.1016/j.ajpc.2020.100010] [Medline: 34327452]
18. Boklage SH, Malangone-Monaco E, Lopez-Gonzalez L, Ding Y, Henriques C, Elassal J. Statin utilization patterns and outcomes for patients with acute coronary syndrome during and following inpatient admissions. Cardiovasc Drugs Ther 2018 Jun;32(3):273-280. [doi: 10.1007/s10557-018-6800-3] [Medline: 29855748]
19. Straka RJ, Taheri R, Cooper SL, Smith JC. Achieving cholesterol target in a managed care organization (ACTION) trial. Pharmacotherapy 2005 Mar;25(3):360-371. [doi: 10.1592/phco.25.3.360.61601] [Medline: 15843283]
20. Gunadi S, Upfield S, Pham ND, Yea J, Schmiedeberg MB, Stahmer GD. Development of a collaborative transitions-of-care program for heart failure patients. Am J Health Syst Pharm 2015 Jul 1;72(13):1147-1152. [doi: 10.2146/ajhp140563] [Medline: 26092965]
21. Başaran Ö, Doğan V, Mert KU, et al. How did the updated 2019 European Society of Cardiology/European Atherosclerosis Society risk categorization for patients with diabetes affect the risk perception and lipid goals? A simulated analysis of real-life data from EPHESUS study. Anatol J Cardiol 2023 Feb;27(2):78-87. [doi: 10.14744/AnatolJCardiol.2022.2012] [Medline: 36747449]
22. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015 Feb 6;350:h391. [doi: 10.1136/bmj.h391] [Medline: 25662947]
23. Wilson EB. Probable inference, the law of succession, and statistical inference. J Am Stat Assoc 1927 Jun;22(158):209-212. [doi: 10.1080/01621459.1927.10502953]
24. Rosenson RS, Farkouh ME, Mefford M, et al. Trends in use of high-intensity statin therapy after myocardial infarction, 2011 to 2014. J Am Coll Cardiol 2017 Jun 6;69(22):2696-2706. [doi: 10.1016/j.jacc.2017.03.585] [Medline: 28571633]
25. Zheutlin AR, Derington CG, Herrick JS, et al. Lipid-lowering therapy use and intensification among united states veterans following myocardial infarction or coronary revascularization between 2015 and 2019. Circ Cardiovasc Qual Outcomes 2022 Dec;15(12):e008861. [doi: 10.1161/CIRCOUTCOMES.121.008861] [Medline: 36252093]
26. Sarak B, Savu A, Kaul P, et al. Lipid testing, lipid-modifying therapy, and PCSK9 (proprotein convertase subtilisin-kexin type 9) inhibitor eligibility in 27 979 patients with incident acute coronary syndrome. Circ Cardiovasc Qual Outcomes 2021 Apr;14(4):e006646. [doi: 10.1161/CIRCOUTCOMES.120.006646] [Medline: 33813856]
27. Wang WT, Hellkamp A, Doll JA, et al. Lipid testing and statin dosing after acute myocardial infarction. J Am Heart Assoc 2018 Jan 25;7(3):e006460. [doi: 10.1161/JAHA.117.006460] [Medline: 29371200]
28. Writing Committee, Lloyd-Jones DM, Morris PB, et al. 2022 ACC expert consensus decision pathway on the role of nonstatin therapies for LDL-cholesterol lowering in the management of atherosclerotic cardiovascular disease risk: a report of the American College of Cardiology Solution Set Oversight Committee. J Am Coll Cardiol 2022 Oct 4;80(14):1366-1418. [doi: 10.1016/j.jacc.2022.07.006] [Medline: 36031461]
29. Partogi M, Gaviria-Valencia S, Aguirre MA, et al. Sociotechnical intervention for improved delivery of preventive cardiovascular care to rural communities: participatory design approach. J Med Internet Res 2022 Aug 22;24(8):e27333. [doi: 10.2196/27333] [Medline: 35994324]


## Abbreviations

**ACS:** acute coronary syndrome
**EHR:** electronic health record
**LDL:** low-density lipoprotein
**PCSK9:** proprotein convertase subtilisin/kexin type 9
**REDCap:** Research Electronic Data Capture

*Edited by N Cahill; submitted 26.03.24; peer-reviewed by AN Ali, J Osborn, L Askin; revised version received 08.02.25; accepted 11.02.25; published 28.03.25.*

*Please cite as:*
*Flo GL, Alzate Aguirre M, Gochanour BR, Hynes KJ, Scott CG, Fink AL, M Arruda-Olson A*
*Pharmacist-Initiated Team-Based Intervention for Optimizing Guideline-Directed Lipid Therapy of Hospitalized Patients With Acute Coronary Syndrome: Pilot Study Using a Stepped-Wedge Cluster Design*
*JMIR Cardio 2025;9:e58837*
*URL: https://cardio.jmir.org/2025/1/e58837*
*doi:10.2196/58837*

# Adherence Patterns of Patients Using Remote Patient Management After Myocardial Infarction: Mixed Methods Persona Approach

Sara M Hondmann[1,2*], MSc; Laura Schrauwen[1*], MSc; Thomas Reijnders[2], PhD; Esmee Stoop[3], MSc; Andrea WM Evers[1,4], PhD; Valentijn T Visch[2], PhD; Douwe E Atsma[2,4,5], PhD; Veronica R Janssen[5], PhD

[1]Health, Medical and Neuropsychology Unit, Institute of Psychology, Leiden University, Wassenaarseweg 52, Leiden, The Netherlands

[2]Department of Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands

[3]Department of Information Technology and Digital Innovation, Leiden University Medical Center, Leiden, The Netherlands

[4]Medical Delta, Leiden University, Delft University of Technology, Erasmus University, Delft, The Netherlands

[5]Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands

[*]these authors contributed equally

**Corresponding Author:**
Sara M Hondmann, MSc
Health, Medical and Neuropsychology Unit, Institute of Psychology, Leiden University, Wassenaarseweg 52, Leiden, The Netherlands

## Abstract

**Background:** Remote patient management (RPM) using smartphone-enabled health monitoring devices (SHMDs) can be an effective, value-added part of cardiovascular care. However, cardiac patients' adherence to RPM is variable. Personas are fictional representations of users with common behaviors, needs, and motivation and can thereby help guide tailoring of interventions to be meaningful and possibly more effective. Personas can be used to understand the needs of the patient group and guide tailoring toward more personalized and effective eHealth intervention.

**Objective:** The aim of this study was to develop data-driven personas for patients with myocardial infarction (MI) based on both quantitative and qualitative results.

**Methods:** This study used a mixed methods design involving (1) database analysis of patients with MI (N=261) SHMD usage data (blood pressure [BP], weight, step count) over the course of a one-year care track and (2) semistructured interviews with patients with MI (N=16) currently using SHMDs. Overall, 12-month adherence rates were calculated based on the number of weeks patients performed the prescribed home measurements with the SHMDs.

**Results:** A cluster analysis was conducted on the self-monitoring data resulting in four distinctive usage patterns labeled as stiff starting (low adherent in first 6 weeks: 13%, 34/261 of users), temporary persisting (decreasing adherence: 24%, 62/261), loyally persisting (continuously adherent: 26%, 68/261), and negligent quitting (nonadherent: 37%, 97/261). Health outcomes (BP, step count, and weight) were analyzed based on these patterns. More adherent usage patterns show better controlled BP when compared to less adherent usage patterns, suggesting that adherence is associated with health outcomes. Patient experiences regarding adherence or nonadherence to the RPM relating to the four distinctive usage patterns were uncovered by means of semistructured interviews, providing insight into adherence factors most relevant for each of the clusters. Thus, 4 distinct personas were developed by data collection (database analysis and semistructured interviews), persona segmentation, and persona creation, named Tamara, Sam, Peter, and Kim.

**Conclusions:** This study identified 4 personas regarding adherence experiences and usage patterns of patients within an RPM care track. Adherent usage patterns were characterized by improved BP and step count. These personas can guide future tailoring of eHealth interventions to maximize patient adherence.

***(JMIR Cardio 2025;9:e56236)*** doi:10.2196/56236

**KEYWORDS**

cardiovascular disease; eHealth; personas; remote patient management; adherence; blood pressure

## Introduction

Cardiovascular disease (CVD) is a leading cause of death and disability worldwide [1]. CVD can partially originate and be perpetuated by several modifiable lifestyle risk factors including an unhealthy diet, insufficient physical activity, and smoking leading to elevated body weight and blood pressure (BP) [2-4]. Addressing these modifiable risk factors can improve symptoms, health outcomes, and overall quality of life [5]. Various patient groups, for example, those who have experienced an ST-elevation myocardial infarction (STEMI) or non-ST acute coronary syndrome (NST-ACS), benefit from tight control of their BP, weight, and physical activity [6]. A promising method to facilitate rehabilitation and tackle lifestyle risk factors is the use of remote patient management (RPM), adjunct to clinic care [7]. RPM is an intervention type that makes use of eHealth methods to deliver care. eHealth comprises a variety of methods to deliver care and promote health using applications, websites, digital health records [8], and smartphone-enabled health monitoring devices (SHMDs). Various studies show that eHealth is as, or more, effective than standard care [9-11]. Close monitoring through RPM can help to reduce hospital admissions, disease progression, risk factors, and allow for early intervention [12,13].

"The Myocardial Infarction (MI) Box" is an RPM intervention for cardiology patients after myocardial infarction in the home context. It makes use of SHMDs to monitor and provide care to patients discharged from the Leiden University Medical Center (LUMC) after a STEMI or NST-ACS event [6,11] Data from the SHMDs are automatically transferred to and integrated into the electronic medical record (EMR) [11], which provides clinicians with more accurate data to make treatment decisions and give patients feedback on important health parameters, such as BP, body weight, physical activity, and heart rhythm for 1 year after discharge [6]. Furthermore, the MI Box is a tool for the stimulation of a healthy lifestyle. Treskes et al [11] demonstrated through a noninferiority trial that The MI Box as stand-alone aftercare treatment is as effective as standard cardiovascular care. However, like other eHealth interventions, it is curtailed by variable adherence. This nonadherence can lead to a loss of guidance for these patients, a worsening of symptoms, and at worst another cardiac event [7,12].

Adherence to the technology within eHealth is variable, including users who adhere long term, those who stop usage after a short period, those who do not use the functionalities as intended, and those who do not use the devices at all [14-16]. Reasons for this spectrum of adherence could be that needs, preferences, and capacities of users are not sufficiently considered [17], leading to a misalignment between the intervention and the users' needs and abilities. Factors influencing (non)adherence in cardiac interventions are varying, including intrapersonal, clinical, health system, and logistical factors [18]. The duration of the intervention possibly influences adherence, with longer interventions having adverse effects on adherence [19]. Tailoring based on theoretical, behavioral, and demographic variables has been associated with more effective interventions [20]. The literature highlights the importance of balancing personalized and generalizable approaches to improve outcomes. In this study, this balance is achieved using data-driven personas, while in other studies (eg, ref iris and foot), this is accomplished through biometric identification systems, which consider both unique individual characteristics and broader applicability to diverse populations [21]. Personas are a method to determine tailoring strategies to correct this misalignment and enhance adherence [22].

Personas are representations of users with common characteristics, behaviors, and needs [23,24] and can be constituted of quantitative and qualitative data [25,26]. Personas can help prioritize problems, direct focus on specific characteristics and needs of subpopulations, highlight and challenge assumptions about populations, and sensitize those creating interventions on how they differ from the patients they serve. Personas are narratives that highlight the user perspective and can thereby guide conversations and changes to the interventions made by health care professionals and designers [22,23,25,27]. For example, personas can guide cocreative sessions by providing a narrative with which stakeholders can discuss issues facing the users and brainstorm design solutions. Research in patients with cancer, diabetes, and heart failure has used personas to gain a better understanding of the population, which allowed for distinguishing behavioral factors influencing acceptance, distinguishing different subpopulations' needs, and influencing implementation [25,28,29]. Although there is not one definitive way to create personas since they are unique to each population and context, there are three general steps in their creation: data collection, persona segmentation, and persona creation [22]. Data collection comprises a combination of both qualitative and quantitative methods, either separately or together. Persona segmentation is creating groups based on similarities in demographic or behavioral variables. Finally, the persona creation step involves the designing of the layout of the persona and what information is included [22]. To our knowledge, personas based on qualitative and quantitative data representing patients using RMP or SHMDs have not yet been published. Personas based on mixed methods can be valuable for they describe the biopsychosocial complexities of adherence [30], thereby providing insight into tailoring methods to improve compliance and ensuring that interventions are patient-centered.

Therefore, the aim of this study was to develop personas of patients using SHMDs in RPM care. We used a mixed methods approach separated into three steps: data collection, persona segmentation, and persona creation. The data collection is further divided into 2 steps: a database analysis of self-management data and generative semistructured interviews. The developed personas can be used for tailoring of RPM interventions, which could lead to increased adherence and improved user experience. Furthermore, this mixed method approach could be generalized to other eHealth interventions striving to understand their target population and enhance adherence.

## Methods

### Materials: The MI Box

The MI Box is an RPM intervention including four SHMDs: a BP monitor (Wireless Blood Pressure Monitor; Withings), a

step counter (Pulse Ox; Withings), a weight scale (Smart Body Scale Analyzer; Withings), and a single-lead electrocariography (ECG) device (Kardia; AliveCor Inc). The devices communicated with the device-dedicated app on the smartphone via Bluetooth [11]. Patients were followed for one year. This patient population has the same diagnosis and all patients have followed a guideline-driven medical therapy (GDMT). This protocol, standardized with regard to interventions, evaluations, and medication, helps control for confounding variables related to treatment variability, isolating the effects of the RPM intervention and its adherence. Patients can immediately see their measurements, and the data is automatically sent to the LUMC and included in the patient dossier [11]. The data are evaluated multiple times per week by the clinician to monitor the health status of each patient. Thus, the therapeutic regimen could be revised based on the results of measurements as well as symptoms. Furthermore, patients consulted with their physician or nurse practitioner 4 times (1, 3, 6, and 12 months after discharge) to discuss rehabilitation, medication, and any other matters of concern. The first and third meetings were video consults, and the second and fourth were outpatient clinic visits. Patients were prescribed to take at least 1 measurement with each device per week.

### Data Collection

The data collection comprised 2 steps. First, self-management data from Box users who have completed the year-long care track was clustered by type and frequency of the measurements taken (BP, steps, and weight). These variables were selected to differentiate the clusters based purely on usage and thereby provide information on adherence of users. Furthermore, type and frequency of measurements were consistently available in the EMR and reliable indicators of use. Health outcomes (BP, weight, and step count) were analyzed based on these clusters. Second, generative interviews were conducted with patients currently using the SHMDs to understand their experience regarding (non)adherence and enrich the found patterns. This integrated analysis enables the creation of personas.

### Database Analysis

The dataset for this analysis was collected from The MI Box database, a module in the EMR where all patients' measurements were continuously stored in real-time. Data were acquired from May 2017 till January 2020 and included data from patients at least one year after discharge, allowing us to find usage patterns on a yearly basis. The obtained longitudinal dataset included a pseudo-ID for each subject, age, gender, the measurement type (BP, weight, and steps), measurement value, and the corresponding timestamp of each measurement. ECG measurements were not available for analysis since these were not saved in the EMR. The dataset was thereby obtained with anonymized data safeguarding participants' privacy.

### Variables

Prior to the analysis, invalid measurements were removed: a low cut-off point of ≥100 steps per day was applied to the pedometer measurements, as this indicated that the device was not worn but only moved around [31]. Subsequently, all measurement data were aggregated from days to frequencies per week within a range of 0 - 54 weeks (duration of the care track with a margin of 2 weeks). Then, the variables were transformed according to the generic minimal instructions communicated by the hospital: using each device at least once a week. Finally, these device-specific variables were summed, resulting in one time series array per subject consisting of 54 variables with values ranging from 0 - 3. A zero indicated that no devices were used, and a three indicated that all three different devices were used in a specific week.

### Cluster Analysis

In order to identify distinct characteristics in heterogeneous samples and cluster them into homogeneous and meaningful groups, a cluster analysis was conducted based on users' usage pattern over time [27,32]. The cluster analysis was based on each user's use of The MI Box over the 1-year care track and not on demographic variables, such as age and sex, or on the clinical measures (CMs). In other words, the resulting clusters were purely based on similarity of usage patterns. A k-means clustering algorithm including the dynamic time warping (DTW) distance measure was used to determine the different clusters. DTW is a distance measure for dynamically comparing time series data when the time indices between comparison data points do not synchronize perfectly [33]. The cluster algorithm was run for k ranging between 2 - 8. To determine the optimal number of clusters, average silhouette scores for the different values of k were calculated. Values approaching 1 indicated that the data point was in the correct cluster, and values approaching -1 indicated that the data point was in the wrong cluster [34]. These clusters were then named and defined into user patterns.

### Usage Pattern Comparisons

Two approaches to compare the usage patterns were conducted. First, the demographic variables per cluster were explored. Second, an explorative ANOVA was performed to examine differences between the clusters regarding their health outcomes (mean Systolic BP [SBP], diastolic BP [DBP], weight, and steps in month 1, 6, and 12). This second analysis was performed with an objective to compare mean values between clusters, which can be of direct clinical relevance. Multiple univariate ANOVAs were used for normally distributed variables, and the nonparametric Kruskal-Wallis test for nonnormally distributed variables. Furthermore, to compare the health outcomes over time within each cluster, multiple repeated measures ANOVAs were executed for the health outcomes in month 1, month 6, and month 12. All analyses were judged at the threshold $P<.05$. To correct for multiple testing, a Bonferroni correction was used.

### Software

The cluster analysis and corresponding data visualization tasks were carried out using Python 3.8 and the following libraries: NumPy for numerical computations [35], Pandas for data structures [36], Plotly for data visualizations [37], Scikit-learn for clustering [33], Tslearn for DTW clustering [38], and Streamlit for data visualizations [39]. Statistical analyses for cluster characterization and exploration were performed in the software package IBM SPSS Statistics 26.

### Generative Interviews

#### *Participants*

A quota sampling strategy for recruiting patients using The MI Box, in collaboration with specialized nurses and cardiologists, was used. Sampling based on usage patterns allowed recruiting a varying group of participants. A decision tree was created to recruit patients based on their usage data. LS and TR both classified participants based on this tree and results were compared until a consensus was reached (refer to Multimedia Appendix 1). Potential participants were approached by the nurse practitioner in the outpatient clinic, and those who expressed interest were contacted by the research team. Once the participants consented, they were formally enrolled and participated in the interviews. To be able to connect a participant to a user profile, participants were required to be in follow-up for at least 6 months. Data of their SHMDs was provided in the EMR.

#### *Procedure*

One week before the interview, each participant received a paper-based sensitizing booklet with four exercises to complete. Sensitizing booklets are part of the context mapping research method [40] and help prepare participants for an interview. Sensitization allows for a greater and high-quality contribution of the participants, as participants will gain insight into their experiences, enabling them to share this during their interview [40]. The sensitizing booklet included questions about their experience with The Box, positive and negative aspects about the use of The Box, experience with the individual devices, possible changes in lifestyle, and possibility to improve The Box. These questions and answers were discussed further in the semistructured interviews. Furthermore, during the interview, participants were asked to reflect on their usage and the found usage patterns. Semistructured interviews were conducted via videocalling with selected patients based on their personal usage pattern. Interviewees provided written informed consent. The interviews lasted between 30 - 60 minutes and were audio-recorded, transcribed, and thematically analyzed using Atlas.ti (by LS and TR) [41,42]. Transcripts were independently coded and then grouped into broader themes, and discrepancies were discussed and resolved.

#### *Persona Development*

The personas were developed based on the clusters stratified and emergent themes from the quantitative and qualitative methods. This integrates both data-driven and qualitative-focused personas development methods and includes 3 steps as described by Alsaadi and Alahmadi (2021): data collection, persona segmentation, and persona creation [22].

### Ethical Considerations

The Medical Research Ethics Committee Leiden The Hague Delft waived ethical approval for this study as the Dutch law concerning research involving human beings (Dutch abbreviation WMO) did not apply to this protocol (protocol N21.048).

## Results

### Database Analysis

#### *Demographics*

In total, 263 subjects were available to be included; however, 2 subjects were excluded as outliers due to one being a test subject and the other having unrealistic measurement frequency. In total, 261 subjects were included in the analyses. The average age was 58 (SD 10.59) years and 77.4% (202/261) were male. Overall, 29% (76/261) of the participants sent at least one measurement each week for 52 weeks and 53% (138/261) of the patients sent data for more than 80% (41.6/52) of the weeks within the care track.

#### *Usage Patterns*

Figure 1 shows the average silhouette scores for k ranging between 2 and 8 was optimal for k=4, yielding an average silhouette score of 0.236, slightly higher compared to the silhouette score of k=3. Therefore, the cluster analysis yielded four distinct usage patterns based on usage over one year; we named these patterns “temporarily persistent (TP),” “stiff starting (SS),” “negligent quitting (NQ),” and “loyally persistent (LP).” Table 1 and Figures 2-5 provide an overview of the usage patterns.

**Figure 1.** Average silhouette scores of K ranging between 2 and 8.

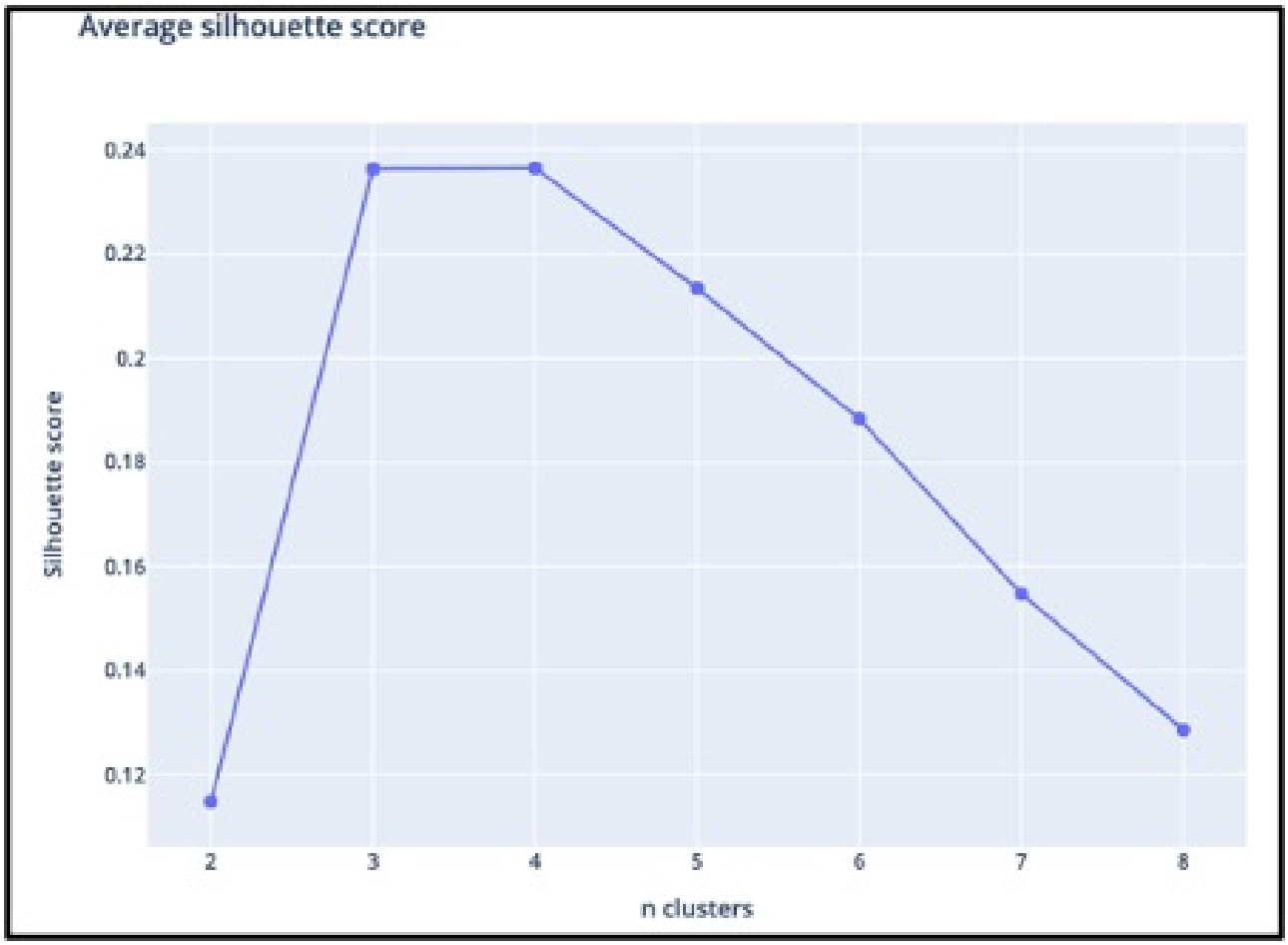



**Table .** Details of the four usage patterns.

| Usage pattern | Demographics (N=261) | Description | Adherence |
|---|---|---|---|
| TP[a] | 62 (24): M[b]: 81% F[c]: 19% | Adherent start with all devices. Adherence decreases as time progresses, possible dropout. | 47% (29/62) measured at least once in 80% (41.6/52) of the weeks. Dropout rate: 19% (12/62) after 6 months and 77% (48/62) after 12. |
| SS[d] | 34 (13): M: 76% F: 24% | Nonadherent. After some weeks, adherent and the adherence declines, possible dropout. | 68% (23/34) measured at least once a week for 80% (41.6/52) of the weeks. Dropout rate: 3% (1/34) after 6 months and 35% (12/34) after 12. |
| NQ[e] | 97 (37): M: 70% F: 30% | Nonadherent start. Adherence decreases over time and often results in a dropout. | 20% (19/97) measured at least once a week in 80% (41.6/52) of the weeks. Dropout rate: 20% (19/97) after 6 months and 66% (64/97) after 12. |
| LP[f] | 68 (26): M: 85% F: 15% | Continuously adherent, with at least two devices. Maintained until the care track cessation. | 100% (68/68) measured at least once in 80% (41.6/52) of the weeks. Dropout rate: 0% (0/68) after both 6 and 12 months. |

[a]TP: temporarily persistent.

[b]M: male.

[c]F: female..

[d]SS: stiff starting

[e]NQ: negligent quitting.

[f]LP: loyally persisting.

**Figure 2.** Measurement pattern temporarily persistent.

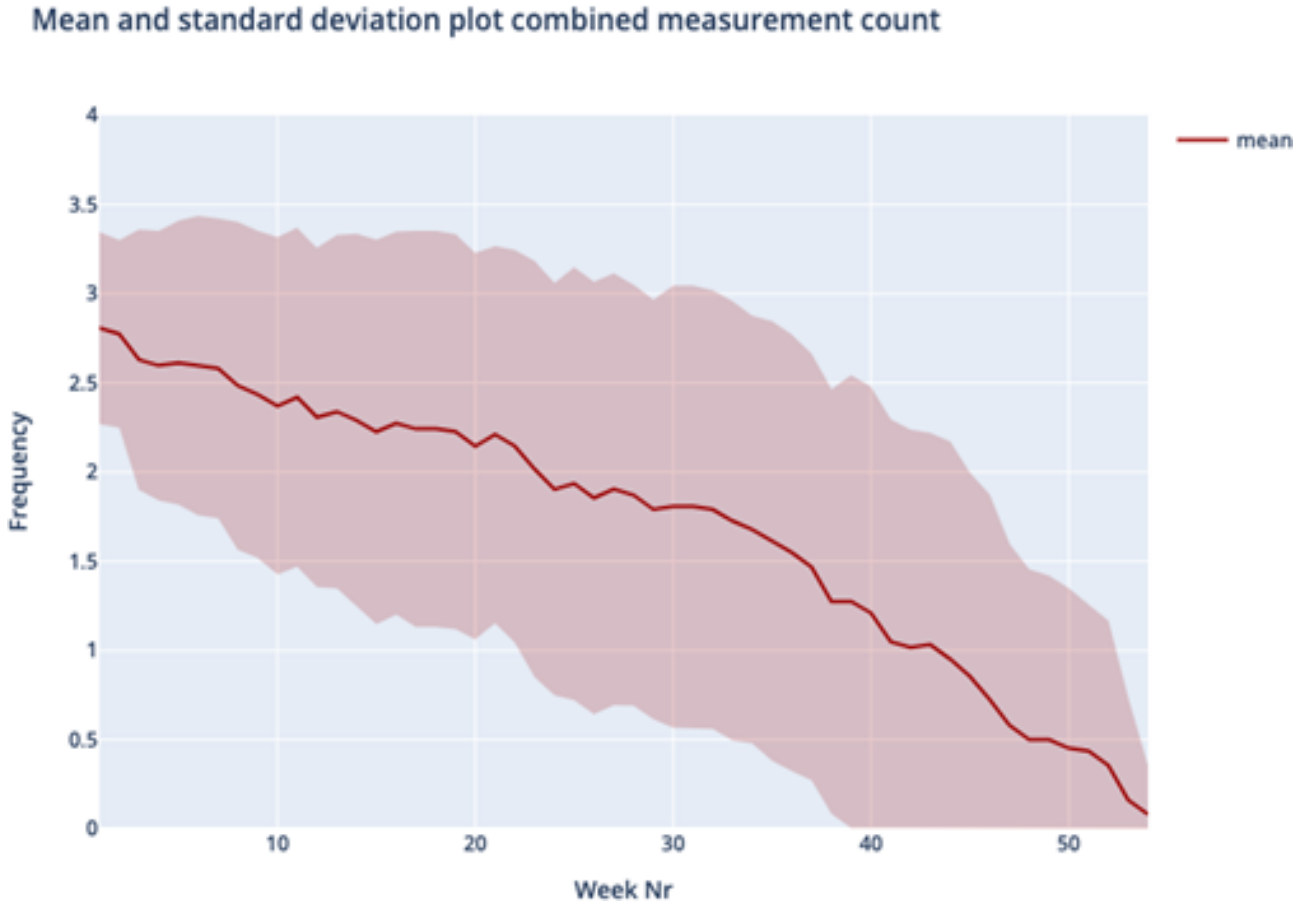



**Figure 3.** Measurement pattern stiff starting.

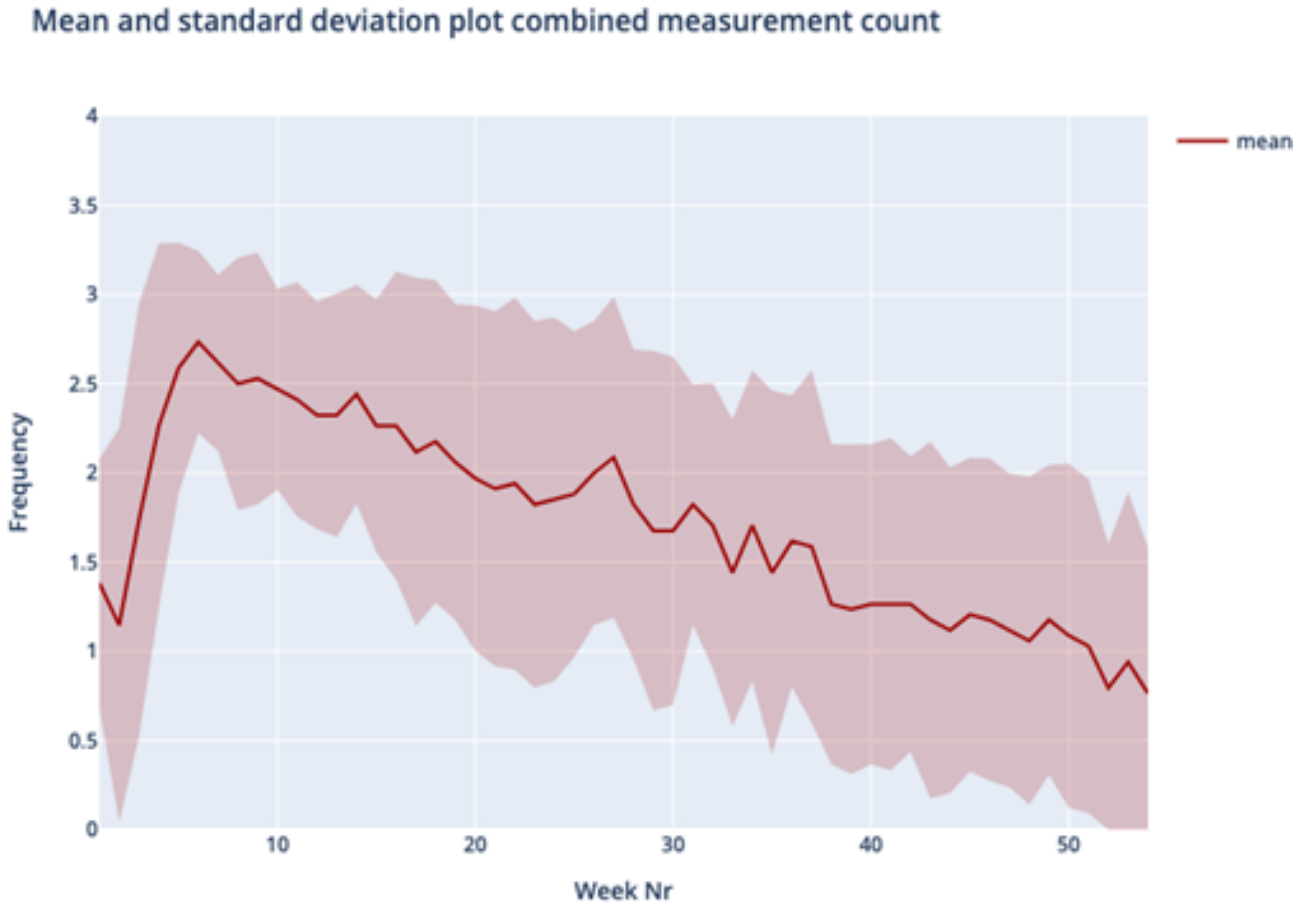


**Figure 4.** Measurement pattern negligent quitting.

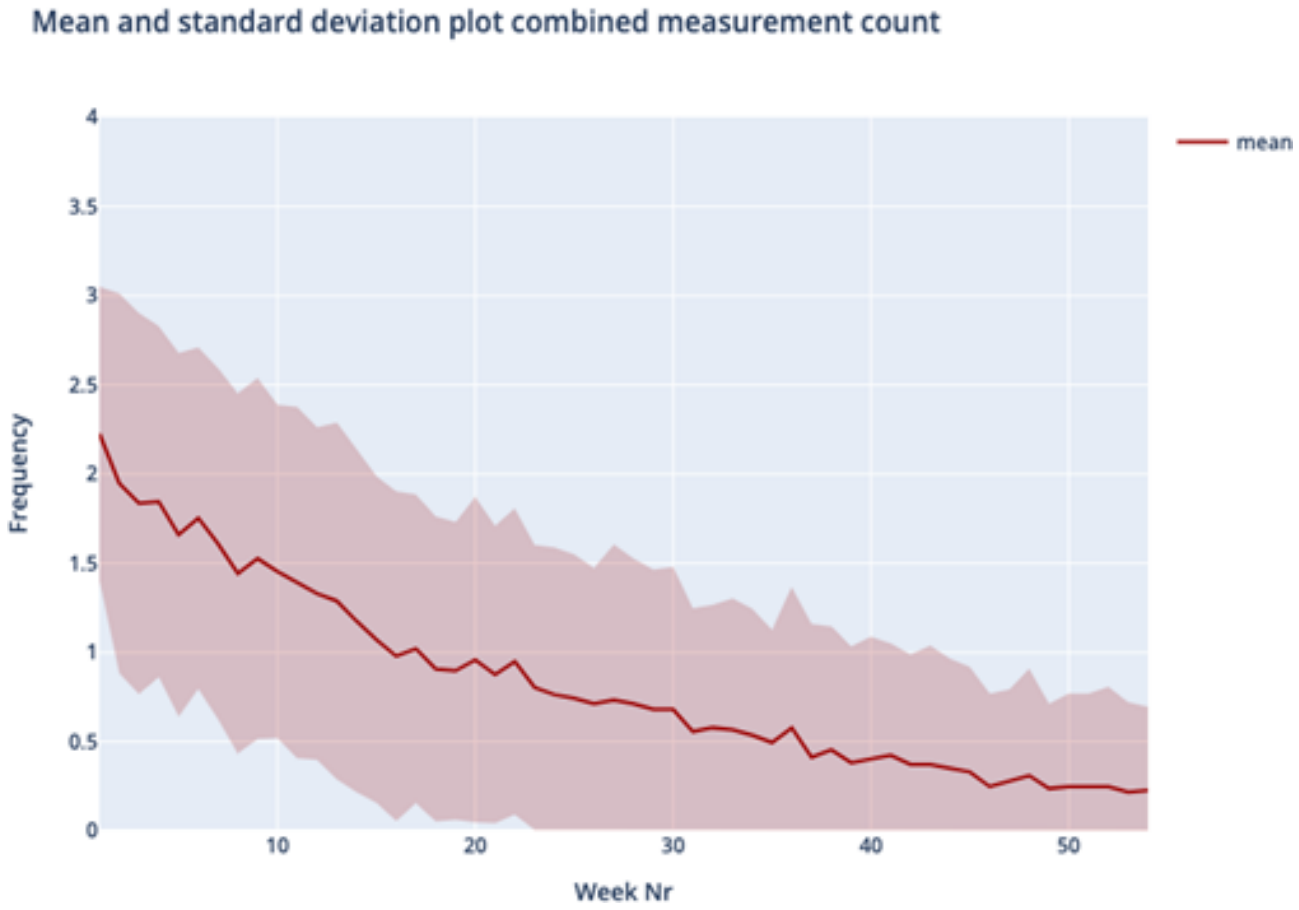



**Figure 5.** Measurement pattern loyally persisting.

### User Pattern Comparisons

The ANOVAs showed a significant difference in DBP between the clusters (NQ and LP and TP and NQ) in month 1 ($F_{3,245}$=4.649, *P*=.004). The ANOVA also showed a significant difference in SBP between the clusters (NQ and LP and TP and NQ) in month 6 ($F_{3,199}$=5.388, *P*=.001). Significance is determined at *P*<.013 (.05/4) after Bonferroni correction for multiple testing.

Additional contrast analysis, comparing the most adherent (LP) with the least adherent (NQ) user profile, showed that patients in the least adherent group had a significantly higher DBP in month 1 and 6 (DBP month 1: $T_{245}$=-3.308, *P*=.001, *r*=0.21, DBP month 6: $T_{199}$=−2.766, *P*=.007, *r*=0.19).

Additionally, a Kruskal-Wallis test was conducted, due to non-normal distribution, on the steps which yielded a significant difference in the number of steps in the first month among the clusters ($\chi^2_3$=9.850, *P*=.020). A higher step count is seen in month one in TP and LP when compared to SS and NQ. The ANOVAs indicated no significant differences in weight between the clusters. Repeated measures ANOVAs did not show significant differences in changes in health outcomes in month 1 - 6, 6 - 12, and 1 - 12 within each cluster. However, especially in month 12, there existed a substantial amount of missing data (see Table 2).

**Table .** Mean values of health outcomes per month, including the percentage available data in brackets.

| CM and month | | TP[a] n (%) | SS[b] n (%) | NQ[c] n (%) | LP[d] n (%) |
|---|---|---|---|---|---|
| Weight | | | | | |
| | 1 | 84.9 (89) | 86.8 (94) | 87.6 (70) | 84.7 (99) |
| | 6 | 84.9 (65) | 85.8 (79) | 86.7 (35) | 84.8 (99) |
| | 12 | 91.2 (21) | 86.3 (38) | 91.8 (14) | 83.4 (91) |
| Steps | | | | | |
| | 1 | 4346 (98)[e] | 3062 (85)[e] | 3559 (66)[e] | 4592 (97)[e] |
| | 6 | 3883 (74) | 4186 (59) | 4331 (23) | 4861 (97) |
| | 12 | 4028 (24) | 5339 (29) | 3885 (5) | 4911 (93) |
| SBP[f] | | | | | |
| | 1 | 124.5 (98) | 127.2 (100) | 127.2 (89) | 124.1 (100) |
| | 6 | 122.9 (74)[e] | 126.5 (97) | 130.7 (58)[e] | 123.5 (100)[e] |
| | 12 | 125.4 (31) | 125.0 (71) | 128.7 (29) | 123.1 (97) |
| DBP[g] | | | | | |
| | 1 | 74.5(98)[e] | 75.6 (100) | 78.5(89)[e] | 74.1(100)[e] |
| | 6 | 74.1 (74) | 75.9 (97) | 78.7 (58) | 74.3 (100) |
| | 12 | 77.3 (31) | 74.5 (71) | 77.6 (29) | 74.2 (97) |

[a]TP: temporarily persistent.

[b]SS: stiff starting.

[c]NQ: negligent quitting.

[d] LP: loyally persisting.

[e]A significant difference between the clusters.

[f]SBP: systolic blood pressure.

[g]DBP: diastolic blood pressure.

### Generative Interviews

#### *Demographics*

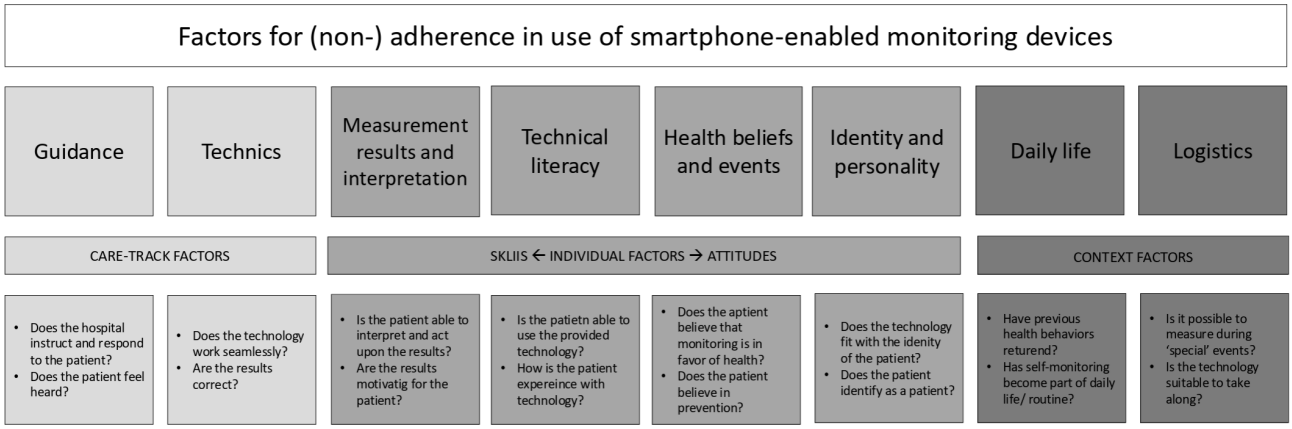
In total, 18 patients were recruited, two canceled their participation, resulting in 16 patients being interviewed. Out of 16, 14 (87.5%) were men and the average age was 60.6 (SD 9.7). Of these patients, 8 (50%) were halfway through the care track and 8 (50%) were approaching the end of the care track. Based on their usage patterns, 6 patients were classified as TP, 2 as SS, 2 as NQ, and 6 as LP (For demographics of interviewed participants, see Multimedia Appendix 2).

#### *Adherence Influencing Factors*

Qualitative analysis in Atlas.ti resulted in 536 codes. Thematic analysis of the codes resulted in eight reasons for adherence, combined into three general factors: care track (guidance and techniques), individual (measurement results and interpretation, technological literacy, health beliefs and events, and identity and personality), and context factors (daily life and logistics).

The questions below the factors are to further explain the meaning of each factor and possible questions to uncover these factors within the patient population (Figure 6).

**Figure 6.** Factors for (non)adherence derived from the interview.




### *Care Track Factors*

#### Guidance

Clear instructions about and response to measurements from the hospital was an important factor for adherence. 11 patients expressed positive experiences with guidance from the hospital: "It is very nice that you are connected to the hospital, they directly called me back the other day!" (TP). In contrast, 5 patients indicated guidance was lacking: "They did not give me any instructions; I just had to see for myself" (SS).

#### Technics

Well-functioning devices lower the threshold for self-management, whereas malfunctioning often resulted in non-adherence. For 9 patients, the devices and app worked seamlessly: "We already had a scale at home, but the provided scale registers everything very nicely" (LP). However, seven patients faced technical problems: "When the devices do not work and are not connecting, I do not know what to do anymore. I have been trying enough for now" (TP).

## Individual Factors

### *Measurement Results and Interpretation*

When patients can interpret the results and consider them meaningful, adherence is more likely. For 12 patients, the outcomes and their interpretations of the measurements were a motivator for measuring: "It is very calming that you see that your blood pressure is stable and your heart rate still good" (LP). In contrast, the measurement results were demotivating for four patients: "If everything was in the red, I thought to myself wat can I do? Should I call the doctor? What can I do?" (TP).

### *Technological Literacy*

The degree of technological literacy and available support can play a critical role in adherence. During the interviews, 11 patients expressed having sufficient experience with technology: "As a physicist, you can call me a 'techie'" (LP). Patients would often receive support from their partner or family members, which made using the devices still doable for them. Five patients indicated a lack of technological literacy: "I never use the app; I am just not able to do that" (NQ).

### *Health Beliefs and Events*

For patients to be intrinsically motivated, they should consider the intervention as relevant for their own health and avoiding future cardiovascular or health events. 11 patients indicated that their infarction and health was a reason for measuring: "When I have health complaints, I just do some extra checks with the devices" (SS). However, five patients did not see the value of measuring for their own health currently and believed that avoiding future events was not greatly influenceable by their own health behavior: "Well, the kilo's do not really matter to me and my health" (TP).

### *Identity and Personality*

It is important that the eHealth intervention aligns with the patient's identity for adherent use. Six patients indicated that the devices matched with their personal identity and identity as a patient. "I have just had an infarction, so it is a fact that I am a patient. I accepted that; it is part of my life now" (LP). On the other hand, 4 patients rejected a patient identity: "I also should not be always labeling myself as a heart patient (SS), or that the smartwatch did not fit their personal style: "For my job, I need to dress appropriately, and the pedometer watch did not fit my style" (SS).

## Context Factors

### *Daily Life*

Whether using the intervention has become part of daily practice can influence adherence. This was the case for 12 patients: "I always measure around 9AM when I just had my breakfast, then I do all the measurements. I also took my medication at that time" (LP). For 4 patients, measuring did not become a habit or was forgotten: "The last weeks I did not do all the measurements, I had a lot of stress at my work. Then I just cannot do it" (LP).

### *Logistics*

It is essential that the telemonitoring devices are easy to handle in varying contexts to stimulate adherent use. Seven patients expressed themselves positively regarding logistics. They valued

that The MI Box reduced their travel time to the hospital and that the devices were easy to take with them: “I do not like to travel a lot, as I always face many obstacles such as tractors, which makes me nervous. Therefore, I really like that I can just do the measurements at a distance” (LP). However, three patients indicated logistical issues which led them not to do their prescribed measurements: “Well, then you have to measure three times when you are on a holiday, I am never going to do that” (LP).

When placing the factors listed in Figure 6 alongside the usage patterns, trends can be observed. Participants in the LP group mentioned that for the majority, all factors contributing to their adherence. Participants in the TP group showed a more divided perspective on factors for adherence. Most participants found guidance, measurement results and interpretation, daily life, and logistics a factor for their adherence. This group is divided on technological literacy and identity and personality. Participants in the NQ group, guidance was a factor for adherence. However, many mentioned health beliefs and measurement results and interpretations to be factors for their nonadherence. Finally, participants in the SS group identified several factors for adherence including technical issues, measurement results and interpretation, technological literacy, health belief, and events. However, guidance was mentioned as the factor for nonadherence. The distribution per usage pattern of the number of participants who mentioned the factor as a reason for adherence and in parentheses those who mentioned it as a factor for nonadherence can be seen in Table 3.

**Table .** Number of participants per usage pattern factor for adherence (and nonadherence in parentheses).

| Factor | LP[a] (N=6) | TP[b] (N=6) | NQ[c] (N=2) | SS[d] (N=2) |
|---|---|---|---|---|
| Guidance | 5 (1) | 4 (2) | 2 (0) | 0 (2) |
| Technical Issues | 5 (1) | 1 (5) | 1 (1) | 2 (0) |
| Measurement results and interpretation | 6 (0) | 4 (2) | 0 (2) | 2 (0) |
| Technological literacy | 5 (1) | 3 (3) | 1 (1) | 2 (0) |
| Health beliefs and events | 6 (0) | 4 (2) | 0 (2) | 2 (0) |
| Identity and personality | 2 (1) | 2 (1) | 0 (1) | 1 (1) |
| Daily life | 6 (0) | 4 (2) | 1 (1) | 1 (1) |
| Logistics | 2(1) | 5 (1) | 0 (1) | 1 (0) |

[a]LP: loyally persisting.

[b]TP: temporarily persistent.

[c]NQ: negligent quitting.

[d]SS: stiff starting.

## Persona Development

Four personas were developed based on the results from the database analysis and generative interviews, named Tamara, Peter, Sam, and Kim. These personas are composed of usage patterns enriched by interview data to provide a rich description of the patient population. Furthermore, each persona contains selected factors from the interviews that are either important or unique within this usage pattern. These personas are not descriptive of all within each usage pattern and are not an identical copy of the data; however, they provide an impression of patients and their concerns with the eHealth intervention (see Multimedia Appendix 3).

### *Tamara*

Tamara represents the loyally persisting usage pattern. The SHMDs motivate her to get and stay healthy. The two factors essential to Tamara are patient identity and routine. Namely, Tamara is okay with being a patient and has built eHealth into her daily routine. Tamara can run the risk of overtesting to keep a sense of control and reduce her feelings of panic about her health. The change in Tamara’s health outcomes (which represents the average LP user) is shown in Table 4.

**Table .** Average change in health outcomes per persona.

| Persona and clinical measure | | Month 1 | Month 6 | Month 12 |
|---|---|---|---|---|
| Tamara | | | | |
| | Weight (kg) | 85 | 85 | 83 |
| | Steps | 4592 | 4861 | 4911 |
| | SBP[a] (mm Hg) | 124 | 124 | 123 |
| | DBP[b] (mm Hg) | 74 | 74 | 74 |
| Peter | | | | |
| | Weight (kg) | 85 | 85 | 91 |
| | Steps | 4346 | 3883 | 4028 |
| | SBP (mm Hg) | 125 | 123 | 125 |
| | DBP (mm Hg) | 75 | 75 | 77 |
| Sam | | | | |
| | Weight (kg) | 85 | 85 | 91 |
| | Steps | 4346 | 3883 | 4028 |
| | SBP (mm Hg) | 125 | 123 | 125 |
| | DBP (mm Hg) | 75 | 75 | 77 |
| Kim | | | | |
| | Weight (kg) | 87 | 86 | 86 |
| | Steps | 3062 | 4186 | 5339 |
| | SBP (mm Hg) | 127 | 127 | 125 |
| | DBP (mm Hg) | 76 | 76 | 75 |

[a]SBP: systolic blood pressure.

[b]DPB: diastolic blood pressure.

### *Peter*

Peter represents the temporarily persisting usage pattern. For Peter, the intervention must be easy, and he does not experience the SHMDs to work seamlessly. The 2 factors essential to Peter are support and technological literacy. Peter has a strong support network of friends and family. However, he is unsure about his technological skills. The change in Peter's health outcomes (average TP user) is shown in Table 4.

### *Sam*

Sam represents the negligently quitting usage pattern. Sam does not see the use of the SHMDs, especially when he is not experiencing any symptoms. The two factors essential to Sam are identity and routine. Sam does not feel like a patient and often forgets to measure. The change in Sam's health outcomes (average NQ user) is shown in Table 4.

### *Kim*

Kim represents the stiff starting usage pattern. At first, Kim did not know all the functions and features of the devices and found it hard to figure out. The two important factors for Kim are technological literacy and support. Kim has confidence in her technological skills, but she does not feel supported by the hospital. The change in Kim's health outcomes (average SS user) is shown in Table 4.

## Discussion

### Principal Findings

The aim of this study was to identify user patterns to improve adherence to RPM interventions in patients after MI. The identified 4 distinct usage patterns are consistent with the literature which found similar usage patterns based on usage data of interventions within different populations, differentiating between those users who do interact, who sparingly interact, and highly interact with technology [15,16]. These usage patterns can help healthcare professionals to identify those patients who are less reached by technology and to identify strategies to improve their adherence to RPM interventions after MI.

Analysis of health outcomes based on these usage pattern clusters showed a significant difference in BP between the patterns (NQ and LP and LP and TP) in month 1 and 6 and for steps in month one (LP and TP and SS and NQ), with a trend of lower blood pressure and higher step count in the high adherence groups. This shows the potential for tailoring to improve health outcomes throughout the whole care track and between different patterns. The differences between usage patterns in months 1 and 6 are possible moments of interest for tailoring to ensure an advantageous start and continuation. Furthermore, around month 6, patients conclude rehabilitation;

therefore, this may be an important moment to address adherence and motivation of users to continue with the intervention. When comparing the most to the least adherent usage patterns, a significantly higher DBP was seen in the least adherent pattern. The observed relation between adherence and better-controlled BP can be explained in several ways. First, medication intake might correlate with device usage [43]. Second, monitoring BP can lead to an increase in healthy lifestyle [44]. Finally, repeated high BP values might demotivate adherently performing the required measurements, as the measurements' results are unsatisfactory [18]. Thus, those with less controlled BP may become less adherent over time. When observing the health outcomes within month one, a possible trend can be seen, namely that users with a "good start" tend to have "good results."

The interviews resulted in care track, individual, and context factors of nonadherence. These findings are consistent with literature [18]. Previous findings have similar aspects for nonadherence, including frustrating technology, perception of content, and support through face-to-face contact [45]. Furthermore, a previous study found similar factors as values of importance for this population and are recommendations for further eHealth development for this population [46].

These findings led to the creation of four personas named Tamara, Kim, Peter, and Sam. These personas are specific to this patient population. However, previous literature within different populations uses similar mixed methods approaches [27,29,47]. Based on these findings, we formulate the following tailoring recommendations. These recommendations were specified per persona; however, these can apply across the care track and for other eHealth interventions. Further research can be conducted on the efficacy of these tailoring measures.

For users such as Peter, the essential factor for adherence is support, and for nonadherence, it is technological literacy. Therefore, comprehensive technology and engagement strategies can be used to tailor the intervention. The more confident a user is in their ability, the more likely they will perform well within the intervention [48]. Furthermore, self-efficacy influences eHealth use and is valued by patients with cardiovascular disease [46,48,49]. Although avoiding all technological difficulties and frustrations is not possible, it is important that users feel that they can reach out for guidance. Therefore, a tailored measure can be to emphasize to these users that technological issues may arise and that this has nothing to do with their competencies; moreover, reaching out to is encouraged. Staff could monitor usage in the first week and be on standby to provide support when nonadherence is observed. Furthermore, these users can be given more time when starting the care track to ask questions and test-run the devices. For users like Kim, whose essential factor for non-adherence is desire for more guidance and support, a recommendation is to provide face-to-face interaction with the health care provider [50]. However, this could be difficult given time constraints of health care providers. Therefore, a recommendation can be to provide the user with a "buddy" or a support group of other participants [51] to which they feel accountable. This buddy can also come directly from the user's social environment [46]. This tailoring measure is essential at the start of the intervention since these users typically start in a non-adherent state. Users like Sam, whose essential factor for nonadherence is identity and routine, can be harder to impact through alterations in the design of the intervention. However, prior to starting the intervention, steps can be made to change their perception and motivation. Tailored, inclusive health education by trusted healthcare professionals or through testimonials from previous users can potentially increase health literacy and their perception of the importance of the intervention [24]. A study on eHealth interventions for smoking cessation showed that those less motivated also engaged less with devices and smartphone applications. Indicating that, regardless of the design elements of the intervention, motivation partially determines engagement. This study indicated that although lack of motivation significantly reduces the engagement of the participant, it is not reduced to zero; therefore, it is critical to find alternative "low-effort communication" [52]. Although relevant to all participants, empowerment strategies could potentially increase these users' adherence. These strategies can include but are not limited to goal setting, feedback on behavior, information about health consequences, social support, and demonstrations of the behavior [48]. Loyal users such as Tamara, whose essential factor for adherence is identity and routine, can provide essential information for improvements to make eHealth more user-friendly. For example, within the interviews, a participant indicated that they forget to do measurements when there is stress in their daily lives. Therefore, providing information on how to continue with measurements when stress is high or how to handle stress could lead to even more consistent use. Furthermore, these users can be supported by further enhancing her autonomy by giving more decision room in the care track [46]. These personas also highlight the potential significance effects of culture and contextual factors on RPM experience and effectiveness; therefore, future research should explore these effects specifically. Focusing on the factors that are crucial to increase adherence and are desirable for a personalized experience.

The study design allowed the obtainment of the aim, to develop personas of patients using SHMDs. Using quantitative insights as the foundation enriched with qualitative insights allowed an in-depth understanding of how SHMDs are used and experienced. The identified usage patterns, adherence factors, and personas are based on objective self-management data and subjective user experience data and thereby provide a deeper understanding of the users and the potential for tailoring. The study approach built upon a by a study of Ten Klooster et al [17], which indicated that meaningful usage patterns can be created through using quantitative data and qualitative insights. Although the study provides a clear stratification and deeper insights, there are limitations worth mentioning. First, the identification of clusters was based on the presence or absence of measurements. Consequently, for nonadherent patients, much data was missing. Therefore, the exploration of CMs was done on the means of month 1, 6, and 12. Second, since no measures were undertaken to predict cluster membership or to apply the results to new data, one should be cautious in generalizing the results beyond the sample. Third, the distribution of user patterns within the interview sample is not equal; however, this was taken into consideration when creating the personas to include

narrative aspects from all patterns. In addition, the interview part of this study involved a relatively small sample size; however, due to the in-depth nature of the interview, we were able to get rich insights into patient experience on adherence. Finally, the male sex made up most of the participants within the study, specifically, the interview study included only 2 females. This could have influenced the creation of the personas, since it has been shown that males and females interact differently with the health care system and providers [53-55].

## Conclusion

The goal of this study was to unlock the personalizing potential of RPM eHealth interventions in motivating and personally meaningful care for patients with MI. The identified usage patterns can help health care professionals to identify those patients who are less reached by technology and to identify strategies to improve their adherence to RPM interventions after MI. This study identified 4 usage patterns and personas, namely temporarily persisting Peter, stiff starting Kim, negligently quitting Sam and loyally persisting with Tamara, provides insights into their reasons for adherence. These personas can assist healthcare professionals in tailoring interventions to the patient subpopulations, aiming at higher adherence and effectiveness of interventions. The usage patterns can indicate whether a patient may drop out, which can result in losing overview of this patient or, as this study suggests, a worsening of BP control. The study provides a deeper understanding of the heterogenous patient population and, to our knowledge, is the first to publish usage patterns and personas for this type of eHealth intervention. Next steps will include using these personas to tailor the RPM intervention to the individual with the aim to improve overall adherence and clinical outcomes.

## Acknowledgments

This study was supported by grants from the Enduring Rewards project. We would like to thank the developers of the LUMCcare application, Innovattic. As well as the members of the BENEFIT consortium. This research was conducted by three different institutions: Leiden University faculty of Health, Medical, and Neuropsychology, Leiden University Medical Center Heart Center, and TU Delft faculty of Industrial Design Engineering.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Decision tree.
[PDF File, 30 KB - cardio_v9i1e56236_app1.pdf ]

Multimedia Appendix 2
Demographics of interviewed participants.
[DOCX File, 16 KB - cardio_v9i1e56236_app2.docx ]

Multimedia Appendix 3
The box personas.
[PDF File, 228 KB - cardio_v9i1e56236_app3.pdf ]

## References

1. Roth GA, Johnson C, Abajobir A, et al. Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. J Am Coll Cardiol 2017 Jul 4;70(1):1-25. [doi: 10.1016/j.jacc.2017.04.052] [Medline: 28527533]
2. Kotseva K, Wood D, De Bacquer D, et al. EUROASPIRE IV: A European Society of Cardiology survey on the lifestyle, risk factor and therapeutic management of coronary patients from 24 European countries. Eur J Prev Cardiol 2016 Apr;23(6):636-648. [doi: 10.1177/2047487315569401] [Medline: 25687109]
3. Kotseva K, Wood D, De Bacquer D, EUROASPIRE investigators. Determinants of participation and risk factor control according to attendance in cardiac rehabilitation programmes in coronary patients in Europe: EUROASPIRE IV survey. Eur J Prev Cardiol 2018 Aug;25(12):1242-1251. [doi: 10.1177/2047487318781359] [Medline: 29873511]
4. Saner H, van der Velde E. eHealth in cardiovascular medicine: A clinical update. Eur J Prev Cardiolog 2016 Oct;23(2_suppl):5-12. [doi: 10.1177/2047487316670256]
5. Janssen V, De Gucht V, Dusseldorp E, Maes S. Lifestyle modification programmes for patients with coronary heart disease: a systematic review and meta-analysis of randomized controlled trials. Eur J Prev Cardiol 2013 Aug;20(4):620-640. [doi: 10.1177/2047487312462824] [Medline: 23022703]
6. Treskes RW, van Winden LA, van Keulen N, et al. Using Smart Technology to Improve Outcomes in Myocardial Infarction Patients: Rationale and Design of a Protocol for a Randomized Controlled Trial, The Box. JMIR Res Protoc 2017 Sep 22;6(9):e186. [doi: 10.2196/resprot.8038] [Medline: 28939546]

7. Verburg A, Selder JL, Schalij MJ, Schuuring MJ, Treskes RW. eHealth to improve patient outcome in rehabilitating myocardial infarction patients. Expert Rev Cardiovasc Ther 2019 Mar;17(3):185-192. [doi: 10.1080/14779072.2019.1580570] [Medline: 30732481]
8. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
9. Bosworth HB, Olsen MK, Grubber JM, et al. Two self-management interventions to improve hypertension control: a randomized trial. Ann Intern Med 2009 Nov 17;151(10):687-695. [doi: 10.7326/0003-4819-151-10-200911170-00148] [Medline: 19920269]
10. McManus RJ, Mant J, Bray EP, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. Lancet 2010 Jul 17;376(9736):163-172. [doi: 10.1016/S0140-6736(10)60964-6] [Medline: 20619448]
11. Treskes RW, van Winden LAM, van Keulen N, et al. Effect of Smartphone-Enabled Health Monitoring Devices vs Regular Follow-up on Blood Pressure Control Among Patients After Myocardial Infarction: A Randomized Clinical Trial. JAMA Netw Open 2020 Apr 1;3(4):e202165. [doi: 10.1001/jamanetworkopen.2020.2165] [Medline: 32297946]
12. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
13. Keesman M, Janssen V, Kemps H, et al. BENEFIT for all: An ecosystem to facilitate sustained healthy living and reduce the burden of cardiovascular disease. Eur J Prev Cardiol 2019 Apr;26(6):606-608. [doi: 10.1177/2047487318816388] [Medline: 30486650]
14. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the Concept of Adherence to eHealth Technology: Systematic Review on When Usage Becomes Adherence. J Med Internet Res 2017 Dec 6;19(12):e402. [doi: 10.2196/jmir.8578] [Medline: 29212630]
15. Tahsin F, Tracy S, Chau E, et al. Exploring the relationship between the usability of a goal-oriented mobile health application and non-usage attrition in patients with multimorbidity: A blended data analysis approach. Digit Health 2021;7:20552076211045579. [doi: 10.1177/20552076211045579] [Medline: 34868614]
16. Wangberg SC, Bergmo TS, Johnsen JAK. Adherence in Internet-based interventions. Patient Prefer Adherence 2008 Feb 2;2:57-65. [Medline: 19920945]
17. Ten Klooster I, Noordzij ML, Kelders SM. Exploring How Professionals Within Agile Health Care Informatics Perceive Visualizations of Log File Analyses: Observational Study Followed by a Focus Group Interview. JMIR Hum Factors 2020 Jan 21;7(1):e14424. [doi: 10.2196/14424] [Medline: 31961325]
18. Resurrección DM, Moreno-Peral P, Gómez-Herranz M, et al. Factors associated with non-participation in and dropout from cardiac rehabilitation programmes: a systematic review of prospective cohort studies. Eur J Cardiovasc Nurs 2019 Jan;18(1):38-47. [doi: 10.1177/1474515118783157] [Medline: 29909641]
19. Jakob R, Harperink S, Rudolf AM, et al. Factors Influencing Adherence to mHealth Apps for Prevention or Management of Noncommunicable Diseases: Systematic Review. J Med Internet Res 2022 May 25;24(5):e35371. [doi: 10.2196/35371] [Medline: 35612886]
20. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed J E Health 2012 Mar;18(2):137-144. [doi: 10.1089/tmj.2011.0062] [Medline: 22381060]
21. Ashraf S, Saleem S, Ahmed T, Aslam Z, Shuaeeb M. Iris and foot based sustainable biometric identification approach. : IEEE Presented at: 2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM) p. 1-6. [doi: 10.23919/SoftCOM50211.2020.9238333]
22. Alsaadi B, Alahmadi D. The use of persona towards human-centered design in health field: review of types and technologies. Presented at: 2021 International Conference on e-Health and Bioengineering (EHB); Iasi, Romania. [doi: 10.1109/EHB52898.2021.9657744]
23. Miaskiewicz T, Kozar KA. Personas and user-centered design: How can personas benefit product design processes? Design Studies 2011 Sep;32(5):417-430. [doi: 10.1016/j.destud.2011.03.003]
24. Vosbergen S, Mulder-Wiggers JMR, Lacroix JP, et al. Using personas to tailor educational messages to the preferences of coronary heart disease patients. J Biomed Inform 2015 Feb;53:100-112. [doi: 10.1016/j.jbi.2014.09.004] [Medline: 25239261]
25. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
26. Pruitt J, Adlin T. The persona lifecycle: keeping people in mind throughout product design. CiNii. URL: http://ci.nii.ac.jp/ncid/BA78191691 [accessed 2025-08-01]
27. Ten Klooster I, Wentzel J, Sieverink F, Linssen G, Wesselink R, van Gemert-Pijnen L. Personas for Better Targeted eHealth Technologies: User-Centered Design Approach. JMIR Hum Factors 2022 Mar 15;9(1):e24172. [doi: 10.2196/24172] [Medline: 35289759]
28. Hall LK, Kunz BF, Davis EV, Dawson RI, Powers RS. The cancer experience map: an approach to including the patient voice in supportive care solutions. J Med Internet Res 2015 May 28;17(5):e132. [doi: 10.2196/jmir.3652] [Medline: 26022846]

29. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: Development of biopsychosocial personas from a study of older adults with heart failure. Int J Med Inform 2017 Dec;108:158-167. [doi: 10.1016/j.ijmedinf.2017.10.006] [Medline: 29132622]
30. Haldane V, Koh JJK, Srivastava A, et al. User Preferences and Persona Design for an mHealth Intervention to Support Adherence to Cardiovascular Disease Medication in Singapore: A Multi-Method Study. JMIR Mhealth Uhealth 2019 May 28;7(5):e10465. [doi: 10.2196/10465] [Medline: 31140445]
31. Thorup C, Hansen J, Grønkjær M, et al. Cardiac Patients' Walking Activity Determined by a Step Counter in Cardiac Telerehabilitation: Data From the Intervention Arm of a Randomized Controlled Trial. J Med Internet Res 2016 Apr 4;18(4):e69. [doi: 10.2196/jmir.5191] [Medline: 27044310]
32. Rapkin BD, Luke DA. Cluster analysis in community research: Epistemology and practice. American J of Comm Psychol 1993 Apr;21(2):247-277. [doi: 10.1007/BF00941623]
33. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12:2825-2830.
34. Shahapure KR, Nicholas C. Cluster quality analysis using silhouette score. Presented at: 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA); sydney, Australia URL: https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=9259989 [accessed 2025-08-01] [doi: 10.1109/DSAA49011.2020.00096]
35. van der Walt S, Colbert SC, Varoquaux G. The NumPy Array: A Structure for Efficient Numerical Computation. Comput Sci Eng 2011;13(2):22-30. [doi: 10.1109/MCSE.2011.37]
36. McKinney W. Pandas: powerful python data analysis toolkit. 2012.
37. Adams C. Learning Python Data Visualization: Packt Publishing; 2014.
38. Tavenard R, Faouzi J, Vandewiele G, et al. Tslearn, A Machine Learning Toolkit for Time Series Data. J Mach Learn Res 2020;21:1-6.
39. Treuille A, Teixeeira T, Kelly A. Streamlit—the fastest way to build and share data apps. 2021.
40. Sleeswijk Visser F, Stappers PJ, van der Lugt R, Sanders EBN. Contextmapping: experiences from practice. CoDesign 2005 Apr;1(2):119-149. [doi: 10.1080/15710880500135987]
41. ATLAS.ti scientific software development gmbh. (year). ATLAS.ti (Version X) [Computer software].: Berlin, Germany: ATLAS.ti Scientific Software Development GmbH URL: https://atlasti.com/ [accessed 2024-10-06]
42. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
43. Fletcher BR, Hartmann-Boyce J, Hinton L, McManus RJ. The Effect of Self-Monitoring of Blood Pressure on Medication Adherence and Lifestyle Factors: A Systematic Review and Meta-Analysis. Am J Hypertens 2015 Oct;28(10):1209-1221. [doi: 10.1093/ajh/hpv008] [Medline: 25725092]
44. Abdullah A, Othman S. The influence of self-owned home blood pressure monitoring (HBPM) on primary care patients with hypertension: a qualitative study. BMC Fam Pract 2011 Dec 30;12(1):143. [doi: 10.1186/1471-2296-12-143] [Medline: 22208768]
45. Lie SS, Karlsen B, Oord ER, Graue M, Oftedal B. Dropout From an eHealth Intervention for Adults With Type 2 Diabetes: A Qualitative Study. J Med Internet Res 2017 May 30;19(5):e187. [doi: 10.2196/jmir.7479] [Medline: 28559223]
46. Bente BE, Wentzel J, Groeneveld RG, et al. Values of Importance to Patients With Cardiovascular Disease as a Foundation for eHealth Design and Evaluation: Mixed Methods Study. JMIR Cardio 2021 Oct 22;5(2):e33252. [doi: 10.2196/33252] [Medline: 34677130]
47. Dekkers T, Hertroijs DFL. Tailored Healthcare: Two Perspectives on the Development and Use of Patient Profiles. Adv Ther 2018 Sep;35(9):1453-1459. [doi: 10.1007/s12325-018-0765-2] [Medline: 30105659]
48. Duff OM, Walsh DM, Furlong BA, O'Connor NE, Moran KA, Woods CB. Behavior Change Techniques in Physical Activity eHealth Interventions for People With Cardiovascular Disease: Systematic Review. J Med Internet Res 2017 Aug 2;19(8):e281. [doi: 10.2196/jmir.7782] [Medline: 28768610]
49. Sol BGM, van der Graaf Y, van Petersen R, Visseren FLJ. The effect of self-efficacy on cardiovascular lifestyle. Eur J Cardiovasc Nurs 2011 Sep;10(3):180-186. [doi: 10.1016/j.ejcnurse.2010.06.005] [Medline: 20674503]
50. Meland E, Maeland JG, Laerum E. The importance of self-efficacy in cardiovascular risk factor change. Scand J Public Health 1999 Mar;27(1):11-17. [doi: 10.1177/14034948990270011001] [Medline: 10847665]
51. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30. [doi: 10.2196/jmir.1602] [Medline: 21393123]
52. Borrelli B, Bartlett YK, Tooley E, Armitage CJ, Wearden A. Prevalence and Frequency of mHealth and eHealth Use Among US and UK Smokers and Differences by Motivation to Quit. J Med Internet Res 2015 Jul 4;17(7):e164. [doi: 10.2196/jmir.4420] [Medline: 26149323]
53. Mo PKH, Malik SH, Coulson NS. Gender differences in computer-mediated communication: a systematic literature review of online health-related support groups. Patient Educ Couns 2009 Apr;75(1):16-24. [doi: 10.1016/j.pec.2008.08.029] [Medline: 19233586]
54. Escoffery C. Gender Similarities and Differences for e-Health Behaviors Among U.S. Adults. Telemed J E Health 2018 May;24(5):335-343. [doi: 10.1089/tmj.2017.0136] [Medline: 28813630]

55. Zhang X, Guo X, Lai KH, Guo F, Li C. Understanding gender differences in m-health adoption: a modified theory of reasoned action model. Telemed J E Health 2014 Jan;20(1):39-46. [doi: 10.1089/tmj.2013.0092] [Medline: 24161004]

## ABBREVIATIONS

**BP:** blood pressure
**CM:** clinical measures
**CVD:** cardiovascular disease
**DBP:** diastolic blood pressure
**DTW:** dynamic time warping
**ECG:** electrocardiography
**EMR:** electronic medical record
**GDMT:** guideline-driven medical therapy
**LP:** loyally persisting
**LUMC:** Leiden University Medical Center
**MI:** myocardial infarction
**NQ:** negligent quitting
**NST-ACS:** non-ST acute coronary syndrome
**RPM:** remote patient management
**SBP:** systolic blood pressure
**SHMD:** smartphone-enabled health monitoring devices
**SS:** stiff starting
**STEMI:** ST-elevation myocardial infarction
**TP:** temporarily persisting

*Edited by N Ainani; submitted 15.01.24; peer-reviewed by J Bierbooms, M Wright, S Farnbach, S Ashraf; revised version received 01.07.24; accepted 06.02.25; published 18.08.25.*

*Please cite as:*
*Hondmann SM, Schrauwen L, Reijnders T, Stoop E, Evers AWM, Visch VT, Atsma DE, Janssen VR*
*Adherence Patterns of Patients Using Remote Patient Management After Myocardial Infarction: Mixed Methods Persona Approach*
*JMIR Cardio 2025;9:e56236*
*URL: https://cardio.jmir.org/2025/1/e56236*
*doi:10.2196/56236*

Original Paper

# Estimating Trends in Cardiovascular Disease Risk for the EXPOSE (Explaining Population Trends in Cardiovascular Risk: A Comparative Analysis of Health Transitions in South Africa and England) Study: Repeated Cross-Sectional Study

Shaun Scholes[1], PhD; Jennifer S Mindell[1], PhD; Mari Toomse-Smith[2], MSc; Annibale Cois[3,4], PhD; Kafui Adjaye-Gbewonyo[5], ScD

[1]Department of Epidemiology and Public Health, University College London, London, United Kingdom

[2]National Centre for Social Research, London, United Kingdom

[3]Burden of Disease Research Unit, South African Medical Research Council, Cape Town, South Africa

[4]Division of Epidemiology and Biostatistics, School of Public Health, University of Cape Town, Cape Town, South Africa

[5]Faculty of Education, Health and Human Sciences, University of Greenwich, London, United Kingdom

**Corresponding Author:**
Shaun Scholes, PhD
Department of Epidemiology and Public Health
University College London
1-19 Torrington Place
London, WC1E 6BT
United Kingdom
Phone: 44 01207 679 1727
Email: s.scholes@ucl.ac.uk

## Abstract

**Background:** Cardiovascular diseases (CVDs) are the leading cause of death globally. Demographic, behavioral, socioeconomic, health care, and psychosocial variables considered risk factors for CVD are routinely measured in population health surveys, providing opportunities to examine health transitions. Studying the drivers of health transitions in countries where multiple burdens of disease persist (eg, South Africa), compared with countries regarded as models of "epidemiologic transition" (eg, England), can provide knowledge on where best to intervene and direct resources to reduce the disease burden.

**Objective:** The EXPOSE (Explaining Population Trends in Cardiovascular Risk: A Comparative Analysis of Health Transitions in South Africa and England) study analyzes microlevel data collected from multiple nationally representative population health surveys conducted in these 2 countries between 1998 and 2017. Creating a harmonized dataset by pooling repeated cross-sectional surveys to model trends in CVD risk is challenging due to changes in aspects such as survey content, question wording, inclusion of boost samples, weighting, measuring equipment, and guidelines for data protection. This study aimed to create a harmonized dataset based on the annual Health Surveys for England to estimate trends in mean predicted 10-year CVD risk (primary outcome) and its individual risk components (secondary outcome).

**Methods:** We compiled a harmonized dataset to estimate trends between 1998 and 2017 in the English adult population, including the primary and secondary outcomes, and potential drivers of those trends. Laboratory- and non–laboratory-based World Health Organization (WHO) and Globorisk algorithms were used to calculate the predicted 10-year total (fatal and nonfatal) CVD risk. Sex-specific estimates of the mean 10-year CVD risk and its components by survey year were calculated, accounting for the complex survey design.

**Results:** Laboratory- and non–laboratory-based 10-year CVD risk scores were calculated for 33,628 and 61,629 participants aged 40 to 74 years, respectively. The absolute predicted 10-year risk of CVD declined significantly on average over the last 2 decades in both sexes (for linear trend; all $P<.001$). In men, the mean of the laboratory-based WHO risk score was 10.1% (SE 0.2%) and 8.4% (SE 0.2%) in 1998 and 2017, respectively; corresponding figures in women were 5.6% (SE 0.1%) and 4.5% (SE 0.1%). In men, the mean of the non–laboratory-based WHO risk score was 9.6% (SE 0.1%) and 8.9% (SE 0.2%) in 1998 and 2017, respectively; corresponding figures in women were 5.8% (SE 0.1%) and 4.8% (SE 0.1%). Predicted CVD risk using the

Globorisk algorithms was lower on average in absolute terms, but the pattern of change was very similar. Trends in the individual risk components showed a complex pattern.

**Conclusions:** Harmonized data from repeated cross-sectional health surveys can be used to quantify the drivers of recent changes in CVD risk at the population level.

***(JMIR Cardio 2025;9:e64893)*** doi:10.2196/64893

**KEYWORDS**

data harmonization; cardiovascular disease; CVD; CVD risk scores; trends; cross-country comparisons; public health; England; South Africa

## Introduction

The global burden of noncommunicable diseases is increasing [1,2]. Cardiovascular diseases (CVDs) in particular lead globally in terms of causes of mortality [3] and often share characteristics with other major noncommunicable diseases. For instance, they tend to increase with age and can be influenced by healthy lifestyle behaviors as well as other demographic, social, and environmental factors. Along with questions on the presence, diagnosis, and treatment of chronic disease–related conditions, population health surveys conducted at regular intervals often include measures of risk factors for CVD, thus providing opportunities to study health transitions.

Understanding the drivers of epidemiological transition in countries that have not followed predicted paths (eg, South Africa) compared with those that have served as examples (eg, England) can provide knowledge on where best to intervene and direct resources to reduce disease burden. The EXPOSE (Explaining Population Trends in Cardiovascular Risk: A Comparative Analysis of Health Transitions in South Africa and England) study uses participant-level data from nationally representative health surveys to examine health transitions by identifying and quantifying the drivers of trends in CVD risk in a middle-income country such as South Africa compared with a high-income nation such as England. Complete details about the EXPOSE study are available in the study protocol [4] and on the study website [5].

To enable empirical investigation of temporal trends in CVD risk, the first phase of the EXPOSE study was to compile harmonized datasets from the national health surveys conducted in South Africa and England [4]. Since 1991, the Health Survey for England (HSE) has monitored the health of the public in England, including regular updates on trends in key indicators such as smoking, physical activity (PA), overweight and obesity, hypertension, diabetes, and self-reported physician-diagnosed CVD [6]. Creating a harmonized dataset from the annual HSE surveys conducted over 2 decades (1998-2017) to model changes in CVD risk over time and decompose its variation (the later phases of the EXPOSE study) was a daunting task due to changes over time in aspects such as survey content, sampling design (inclusion of boost samples for population subgroups), question wording (eg, through changes in public health policy recommendations), introduction of nonresponse weighting, changes in measuring equipment (eg, changes in blood pressure [BP] monitors), and more stringent data release guidelines for protecting participant anonymity.

Herein, we describe the methods and procedures used to painstakingly compile the harmonized dataset for England, enabling the modeling of trends in CVD risk in adults and the investigation of the factors driving the trends. We anticipate that the dataset will be a valuable resource for the wider research community in the United Kingdom and worldwide (eg, by avoiding duplication of effort). The code for harmonizing and appending the England surveys for others to use in future research is publicly available through the study website [5] and from DataFirst [7]. For the presentation of early results, we provide sex-specific estimates of the mean total (fatal and nonfatal) 10-year CVD risk and its individual risk components (eg, BP, smoking, and physician-diagnosed diabetes) by survey year over 2 decades (1998-2017), accounting for the complex survey design.

## Methods

### The HSE

Data for England were drawn from the HSE, conducted from 1998 to 2017. The HSE is an annual cross-sectional, general population survey of individuals living in private households, with a new sample of addresses selected each year using random multistage stratified probability sampling. Complete details about the HSE, including its origins, sampling design, study content, and data availability, are provided in the "Cohort Profile: The Health Survey for England" [6].

Data collection for each survey was conducted continuously throughout the year, starting in January, to minimize seasonal effects. The process was carried out in 2 stages. The first stage was a computer-assisted health interview, including questions about sociodemographic factors, diagnosed health conditions, self-rated general health and illness, health-related lifestyle behaviors, and direct measurements of height and weight, by trained interviewers. The second stage was a nurse visit, including questions regarding current use of prescribed medicines, BP and other anthropometric measurements (eg, waist and hip circumference), and collection of nonfasting blood samples (eg, glycated hemoglobin [$HbA_{1c}$] and cholesterol). Only those participants who completed the interview were eligible for the nurse visit. Interviews and nurse visits took place in the participants' home. All adults (maximum 10) in selected households were eligible to take part; the percentage of eligible households participating ranged from 74% in 1998 to 59% in 2016.

The survey usually focuses on multiple health issues. The inclusion of a set of "core" questions and measurements each

year (or repeated at regular intervals) provides consistency that is important for studying temporal trends in key health indicators. Some surveys included a greater focus on different single health topics, including PA and fitness in 2008 [8] and respiratory health in 2010 [9]. In a number of years, sampling was boosted to study specific subgroups of the population, including ethnic minority groups in 1999 [10] and 2004 [11], persons living in care homes in 2000 [12], children and young adults in 2002 [13], and persons aged ≥65 years living in private households in 2005 [14]. During these years, a smaller sample of the general population was also selected, with reduced survey content typically limited to the core set of questions and measurements (height and weight).

Through the combination of a health interview and health examination, data from the HSE can be used to investigate both diagnosed and undiagnosed disease at a point in time; a key strength therefore is that each sample is not selected based on health care use [15].

## Ethical Considerations

Each selected address for the HSE receives an advance letter introducing the survey and informing recipients that an interviewer will be visiting to request permission for an interview. Individual interviews are conducted with adults who give verbal informed consent. At the end of individual interviews, participants are asked for agreement to a follow-up visit by a trained nurse. Written consent is obtained for collection of nonfasting blood samples. The advance letters and information leaflets clearly state that participation in the survey is voluntary. Participants are also informed that they may choose not to answer specific questions, withdraw or stop at any time, or refuse any particular measurement if they wish. Interviewers and nurses will often repeat this information in their introductions, when they are setting up appointments, and throughout the interview as necessary. In fact, many individuals choose not to participate in the survey. Others may refuse to answer specific questions, discontinue the interview midway, or decline physical measurements. It is also standard practice to conduct interviews and nurse visits sometime after an appointment has been made so that individuals have a chance to reflect on their agreement before the appointment takes place. The procedures used in the HSE to obtain informed consent are very closely scrutinized by a National Health Service ethics committee each year (complete details are available in the annual HSE "Methods and documentation" reports). Information leaflets and both the content and wording of questionnaires are also carefully reviewed by the ethics committees.

The original data collection was approved each year by a National Health Service research ethics committee. The present analysis did not receive approval from a research ethics committee. The secondary analysis did not need ethical approval, as we used publicly available datasets [16-33]. The authors had permission to use the data.

## Creating a Harmonized Dataset

### *Selection of Participants for Inclusion*

In the survey years including minority ethnic boost samples (1999 and 2004), nurse visits were offered to participants in the target minority ethnic groups only. As systolic BP (SBP, a component of cardiovascular risk scores) was measured during the nurse visit, the harmonized dataset does not include data from the 1999 and 2004 surveys. In addition, we excluded data from 2000 as the question on diagnosed diabetes was not included (also a component of CVD risk), and we included only those participants selected as part of the general population sample in the boost year of 2002. Taken together, the datasets covered 17 cross-sections of the adult population spanning the 20-year period from 1998 to 2017: these datasets are available to registered users via the UK Data Service and were compiled and appended to create the harmonized dataset [16-33].

### *CVD Risk Algorithms*

#### Overview

##### *Background*

The predicted 10-year cardiovascular risk for HSE participants was calculated using laboratory-based and non–laboratory-based algorithms. Risk algorithms such as the Framingham Risk Score and those developed in England and Wales using the QResearch database are widely used in clinical and other settings to predict the risk of a future CVD event based on a number of laboratory results (eg, blood samples) and other demographic and self-reported risk factors [34]. Non–laboratory-based algorithms, based on physical examination and self-reported data, were developed for use in low-resource environments where laboratory-based measures may be difficult to obtain. In this study, we selected the World Health Organization (WHO) [35] and Globorisk [36,37] CVD risk algorithms for several reasons. Both are "global" models, accounting for differences in levels of CVD risk factors and event rates across populations, making them applicable to low-, middle-, and high-income countries. Both algorithms include the "traditional" CVD risk factors—age, sex, SBP, current smoking, diabetes, total cholesterol, and BMI—that are available in both the HSE and in South African datasets such as the Demographic and Health Surveys and the South Africa National Health and Nutrition Examination Survey, thereby fitting in line with the objective of comparing health transitions (using CVD risk as a case study) in these 2 countries. Finally, the statistical code for both algorithms is openly accessible to calculate the predicted 10-year CVD risk for participants in health surveys such as the HSE.

Both algorithms calculate the predicted 10-year risk of CVD, expressed as a proportion or a percentage, based on (1) an individual's risk factor profile (eg, age, current smoking status, BP, total cholesterol, and diabetes history) and (2) the average CVD risk in the target population based on population levels of risk factors (obtained from national health surveys) and rates of CVD. Model derivation and recalibration were performed in both approaches in a broadly similar fashion. At the model derivation stage, individual-level data from prospective cohort studies were used to estimate hazard ratios (HRs) for each risk factor; these quantify the proportional effect of risk factors on CVD risk over the follow-up period. At the model recalibration stage, average risk factor levels and annual CVD event rates were reset to the levels observed in the target population to bring predicted risks in line with observed risks [37].

### *WHO Risk Score*

The WHO algorithm predicts 10-year risk for the combined outcome of fatal and nonfatal CVD based on the revised WHO CVD risk models that have been recalibrated to reflect the expected 10-year risk in contemporary populations in 21 Global Burden of Disease (GBD) regions [35].

Risk prediction models were derived using individual participant data (aged 40-80 years with no baseline CVD) from 85 prospective cohorts mostly from high-income countries in the Emerging Risk Factors Collaboration. Follow-up was until the first CVD event; outcomes were censored if cases were lost to follow-up, died from non-CVD causes, or reached 10 years of follow-up. Variables were considered for inclusion in the risk models if they were known to predict CVD in diverse populations, were available in recent national health surveys for model recalibration within GBD regions, and could be measured at a low cost in low- and middle-income countries.

A laboratory-based CVD model included age, current smoking status, SBP, diabetes history, and total cholesterol; a non–laboratory-based model replaced diabetes and total cholesterol with BMI. Sex-specific models were fitted separately for (1) coronary heart disease (CHD; fatal-plus-nonfatal myocardial infarction or CHD death) and (2) fatal-plus-nonfatal stroke outcomes to enable separate recalibration before combination in a single equation for CVD [35]. HRs were estimated using Cox proportional hazards models, stratified by study and with duration (time-in-study) as the time scale. Interaction terms allowed the proportional effects of risk factors on the risk of CVD to vary by age (as evidence suggests that their impact declines with age).

Models were then recalibrated to the contemporary circumstances of the 21 GBD regions. The recalibration process is broadly similar for the WHO and Globorisk algorithms and involves resetting the average levels of risk factors and CVD risk to the levels observed in the target population. The input data and the steps involved in the model recalibration process, drawing largely on the worked example by the Cohorts Consortium of Latin America and the Caribbean [38], are described as follows.

Input data for model recalibration comprises (1) an individual's risk factor profile (eg, age, sex, SBP, and current smoking status); (2) region-, sex-, and age-specific mean risk factor levels (eg, mean SBP and prevalence of current smoking); and (3) region-, sex-, and age-specific annual rates of CVD events. For the WHO algorithm, region-specific risk factor values were estimated by averaging country-specific levels provided by the Noncommunicable Disease Risk Factor Collaboration [39-43]; CVD incidence rates were obtained from the 2017 update of the GBD study [44,45].

The following steps in the model recalibration process refer to calculations performed separately for each year of follow-up over a period of 10 years (year 0 to year 9). First, for each risk factor, the difference ("distance") is calculated between an individual's risk factor profile and the group-specific mean risk factor levels. Second, for each risk factor, the distance is multiplied by the main coefficient (log HR) of the corresponding risk factor from the relevant (outcome-specific) Cox regression model. Third, for the risk factors whose proportional effect on the outcome varies by age, the distance (eg, individual SBP minus population mean SBP) is multiplied by the coefficient (log HR) of the interaction term and by the individual's age (eg, for someone aged 60 years at year 0 through to age 69 years at year 9). Fourth, for each risk factor, the products obtained from steps 2 and 3 are summed and then exponentiated to calculate the risk factor–specific HR. Fifth, the risk-factor specific HRs are multiplied to compute the joint HR. Sixth, the 1-year risk of CVD is calculated as the product of the joint HR and the group-specific annual CVD event rate. Seventh, the 1-year survival is calculated as the exponential of the negative value of the 1-year risk of CVD (eg, a 1-year CVD risk of 0.06 translates to a 1-year survival of exp(–0.06)=0.942).

In the eighth stage, the cumulative survival is calculated as the product of the 1-year survival in year T and the survival in year T–1. In the ninth and final stage, the cumulative CVD risk is calculated as 1 minus the cumulative survival.

The cumulative CVD risk in the final year of follow-up (year 9) is the predicted *absolute* 10-year CVD risk. For example, based on a survey participants' risk factor profile, a CVD risk of 9% can be interpreted as slightly less than a 1 in 10 chance of having a CVD event in the next 10 years. To facilitate interpretation, CVD risk scores are often categorized into groups such as "very low" (<5%), "low" (5%-10%), "moderate" (10%-20%), "high" (20%-30%), and "very high" (≥30%), and these cutoffs are often used in applications to estimate the proportion of individuals at high absolute CVD risk.

The individual risk factor components of the WHO CVD risk scores and the HSE survey years available for the calculation of CVD risk scores are summarized in Textbox 1. Laboratory-based WHO CVD risk scores are calculated using complete risk factor profile data on sex, age, current smoking status, SBP, history of diabetes, and total cholesterol. (To be comparable with South African data, diabetes status in this study was defined using only self-reported physician-diagnosed diabetes). The non–laboratory-based risk score replaces diabetes and total cholesterol with BMI.

Calculation of CVD risk in our study was limited to participants aged 40 to 74 years. Data on all components of the laboratory-based risk score were available in 1998, 2003, 2006, and from 2009 onward; all components of the non–laboratory-based score were available in 1998, 2001 to 2003, and from 2005 onward. In 2006, participants aged ≥65 years were allocated at random to either (1) the CVD (including diabetes) and short PA modules or (2) the long PA module but not the CVD module. Adults aged 16-64 years completed both the CVD and long PA modules. Herein, for the presentation of CVD trends, components were set to missing for a small number of participants with the following outlying values: *SBP* (<60 mm Hg or >270 mm Hg), *height* (<1.2 m or >2.2 m), *weight* (men: <35 kg or >250 kg; women: <25 kg or >250 kg), *BMI* (<10 kg/m$^2$), and *total cholesterol* (<1.8 mmol/L or >20 mmol/L).

Total (ie, fatal and nonfatal) CVD risk scores for participants with valid data on all the relevant components (ie, complete cases) were calculated using the Stata (version 18.0; StataCorp) program *whocvdrisk*. A 10-year risk time was specified, with Great Britain as the country code identifier (included in the Western European GBD region) and the 2017 update of the GBD study as the base for recalibration parameters.

**Textbox 1.** World Health Organization cardiovascular disease (CVD) risk scores calculated using Health Survey for England data.

**Laboratory based (1998, 2003, 2006, and 2009-2017)**

- Age (40-74 y)
- Sex
- Systolic blood pressure (SBP)
- Physician-diagnosed diabetes
- Current smoking
- Total cholesterol

**Non–laboratory-based (1998, 2001-2003, and 2005-2017)**

- Age (40-74 y)
- Sex
- SBP
- Current smoking
- BMI

### *Globorisk Score*

The Globorisk algorithm calculates the predicted 10-year risk of CVD (CHD or stroke).

Risk prediction models were derived using individual participant data (aged ≥40 years with no baseline CVD, with a maximum follow-up of 15 years) pooled from 8 prospective United States–based cohorts. Cohort-specific models were developed for (1) fatal CVD and (2) fatal-plus-nonfatal CVD (for countries with available data on CVD incidence) using the same set of risk factors as described in the WHO Risk Score section. HRs were estimated using Cox proportional hazards models, including interaction terms to allow for age and sex differences in the effects of risk factors on CVD risk (eg, the estimated associations of diabetes and smoking were observed to be stronger in women) [36,37].

Using a similar process as described in the WHO Risk Score section, models were then recalibrated by applying the risk equation to national-level data on risk factor levels and CVD event rates to calculate the predicted 10-year CVD risk.

The laboratory-based Globorisk score calculated the predicted 10-year risk of CVD in adults aged 40 to 74 years using age, sex, SBP, diabetes (based on blood sugar levels or having a history of diabetes), smoking status, and total cholesterol [36,37]. The prediction was limited to those aged 40 to 74 years, as this age range is commonly used for assessment of primary prevention of CVD. The non–laboratory-based score replaces diabetes and total cholesterol with BMI. Globorisk scores are contemporarily recalibrated for the target country [36-38]; for our study, we specified the population of Great Britain and the baseline year of 2020 and calculated the risk scores for fatal-plus-nonfatal CVD. Globorisk scores for HSE participants were computed using the same analytical samples and risk factor definitions as for the WHO algorithms and were calculated using the R (version 4.2.2; R Foundation for Statistical Computing) package *Globorisk* [46].

## CVD Risk Score Components

### *Age*

All adults (defined as aged ≥16 years in the HSE series) selected in the general population sample in the relevant survey years, who completed the health interview, were included in the harmonized dataset. Since 2015, only categorical age (16-17 years, 18-19 years, and in 5-year intervals up to age ≥90 years) has been provided in the end-user license (EUL) datasets to preserve anonymity of participants. Continuous age (up to ≥90 years) was provided in the special license (SL) dataset for 2015 (SL data collections contain more detailed information than EUL data). For participants in the HSE 2016-2017, age in our study was set to the midpoint of categorical age (data under the 2016-2017 SL was not available at the time of writing this manuscript).

### *Cigarette Smoking Status*

Participants were asked whether they had ever smoked a cigarette, and those who reported having ever smoked were asked whether they smoked cigarettes at all nowadays. Participants aged ≥25 years were asked about their smoking behavior during the interview. In the HSE series, participants are classified as current smokers, ex-smokers, or never smokers. A binary smoking variable (current smoker or not current smoker) was used in our study to calculate CVD risk.

### *Calculation of BMI*

BMI data are derived from measured height and weight. Toward the end of the interview, height was measured by trained interviewers using a portable stadiometer with a sliding head

plate, a base plate, and connecting rods marked with a measuring scale. Participants were asked to remove their shoes. One measurement (to the nearest even millimeter) was taken, with participants stretching to the maximum height and the head positioned in the Frankfort plane. For participants who were not pregnant, a single weight measurement (to the nearest 100 g) was recorded using digital scales. Participants were asked to remove their shoes and any bulky clothing or heavy items from their pockets. Individuals who were unable to stand or were unsteady on their feet were not measured. The participants who weighed >130 kg (>200 kg since 2011) were asked for their estimated weight due to concerns about the accuracy of the scales above these levels. (Class III Seca scales were introduced in the HSE 2011; these met a higher specification than previous [class IV] scales and measure up to a maximum of 200 kg.) Participants were assigned missing values if they were considered by the interviewer to have unreliable measurements, for example, those who were too stooped or wore excessive clothing. Height and weight measurements were voluntary. A sizeable and increasing number of participants had missing anthropometric data; our own analyses of HSE 2003-2018 data showed that the propensity to have missing values was associated with older age, lower educational status, and fair, bad, or very bad general health [47]. BMI was calculated as weight in kilograms divided by height in meters squared, and the WHO obesity classification was used to group participants into mutually exclusive categories [48].

#### *SBP Measurement*

BP was measured during the nurse visit using standardized protocols; Dinamap (Critikon) 8100 monitors were used before 2003, and Omron (Omron Healthcare Co Ltd) HEM 907 have been used since. Dinamap readings were converted into Omron readings using a regression equation based on a calibration study [49]. Three BP readings were taken from each participant while seated, at 1-minute intervals, using an appropriately sized cuff on the right arm, if possible, after a 5-minute rest. Measurements from participants who had exercised, eaten, drunk alcohol, or smoked in the 30 minutes before measurements were recorded as not valid. The mean of the second and third valid SBP readings was used in our study.

#### *Treatment for High BP*

Use of antihypertensive medication is a component of the Framingham Risk Scores [34]. Nurses recorded the details of any classes of medication for high BP that participants reported taking at the time of the survey. Since 2003, participants taking medicines that lower BP were asked whether they were taking the medicine because of a heart problem, high BP, or for some other reason. Two different definitions of use of BP medicine are therefore available [50]. First, participants can be classified as being on treatment if the BP medicine they were taking was prescribed specifically to treat their BP. Second, participants can be classified as being on treatment if they were taking any medicines commonly used to treat high BP, regardless of whether the medicines were reported by the participant as being prescribed for that reason. The former (more restrictive) definition has been used in the HSE series from 2003 onward to classify participants as having survey-defined hypertension (ie, SBP ≥140 mm Hg or diastolic BP ≥90 mm Hg or taking medicine prescribed for high BP) [51].

#### *Diabetes*

The item on physician-diagnosed diabetes was included in the main interview in 1998, 2003, 2006 (all adults aged 16-64 years, but a random half of those aged ≥65 years), and each year from 2009 onward. The interview made no distinction between type 1 and type 2 diabetes. In addition, $HbA_{1c}$ levels were measured in nonfasting blood samples collected at the nurse visit. $HbA_{1c}$ reflects average blood sugar levels over the previous 2 to 3 months and can therefore be used both to monitor diabetic control in people with diagnosed diabetes and to detect undiagnosed diabetes [52]. In the HSE series, $HbA_{1c}$ values expressed as a percentage were available in 2003, 2005 to 2006, and from 2008 onward; $HbA_{1c}$ levels reported in SI units of mmol/mol were available from 2012 onward. The latter is currently used in the annual HSE Adult Health reports to define total diabetes, which is characterized by an $HbA_{1c}$ level of ≥48 mmol/mol (diagnostic of diabetes) or self-reported diagnosed diabetes [53]. Due to changes in calibrators, $HbA_{1c}$ values were adjusted upward from the fourth quarter of fieldwork for the HSE 2013 onward to ensure comparability with earlier years. In our analyses (not presented herein), $HbA_{1c}$ values expressed as a percentage (nonoutlying values: between 2.5% and 24.9%) were converted to mmol/mol values using a conversion equation [54].

#### *Total Cholesterol*

Cholesterol levels were measured via nonfasting blood samples taken at the nurse visit. Due to a change in calibrators, cholesterol levels between 2011 and 2014 were adjusted downward to ensure comparability with values from earlier years. A further change in calibrators in 2015 resulted in equivalence between the measurements in current years and those before 2010.

### Harmonized Variables to Adjust for Change in Measuring Equipment

To avoid duplication of effort, we have provided variables in the harmonized dataset that researchers can use to suitably adjust for the changes over time in the machinery used in the HSE to measure BP, total cholesterol, and $HbA_{1c}$. These are shown in Table 1.

**Table 1.** Harmonized variables to adjust for changes in measuring equipment.

| CVD[a] risk factor | Adjustments | Harmonized variable |
|---|---|---|
| **BP[b,c]** | | |
| Systolic BP | 8.90 + (Dinamap × 0.91) | omsysval |
| Diastolic BP | 19.78 + (Dinamap × 0.73) | omdiaval |
| **Total cholesterol[d]** | Unadjusted minus 0.1 mmol/L | cholval13 |
| **$HbA_{1c}$[e] (mmol/mol)[f]** | | |
| Lower range | 16-41: +1 mmol/mol | glyhb2_h |
| Middle range | 42-68: +2 mmol/mol | glyhb2_h |
| Higher range | ≥69: +3 mmol/mol | glyhb2_h |

[a]CVD: cardiovascular disease.

[b]BP: blood pressure.

[c]Blood pressure was measured using standardized protocols with the use of Dinamap (Critikon) 8100 monitors before 2003 and Omron (Omron Healthcare Co Ltd) HEM 907 from 2003 onward. In the creation of the harmonized dataset, the pre-2003 Dinamap values were converted to Omron values using previously published regression equations based on a calibration study that derived predicted Omron readings from the observed Dinamap readings [49].

[d]New analytical equipment was introduced in April 2010 and June 2015 by the laboratory that carried out the analyses on the blood samples taken during the nurse visit, which resulted in a slight change in the reference range for total cholesterol. For the harmonized dataset, the laboratory values were adjusted downward by 0.1 mmol/L to be comparable to the values before April 2010. For the new equipment introduced post 2015, the laboratory values were on average 0.1 mmol/L lower than the equipment used between 2010 and 2015; hence, no adjustment was needed to be comparable to the values before April 2010 [55].

[e]$HbA_{1c}$: glycated hemoglobin.

[f]A new calibration lot for the processing of glycated hemoglobin was introduced in September 2013. Comparisons by the manufacturer indicated that the new machinery produced lower values, necessitating upward adjustment to be comparable with values before the change in equipment [55].

## Explanatory Variables for Changes in CVD Risk Over Time

### *Socioeconomic Status*

Measures of individual-level socioeconomic status (SES) included educational status, social class, and household income. Educational status was classified into 4 categories according to the highest educational qualification: (1) university degree or equivalent, (2) A level or diploma, (3) O level, General Certificate of Secondary Education, or vocational equivalent, and (4) none. The occupational (social) class was determined using the registrar-general's classification (professional, managerial technical, skilled nonmanual, skilled manual, semiskilled manual, unskilled manual, unemployed, and other or not fully described). The household reference person reported annual gross household income from all sources via a showcard with 31 income categories. Household income was equivalized by considering the number of adults and dependent children in the household (McClements scale [56]); households were divided into quintiles. Tenure, availability of a car, and number of cars normally available for use by household members are also included as other measures of individual-level SES.

Area-level SES was classified in the HSE datasets (from 2001 onward) according to the index of multiple deprivation (IMD). This is a composite index of relative deprivation at lower-layer super output area (LSOA) level, based on 7 domains of deprivation: (1) income, (2) employment, (3) health deprivation and disability, (4) education, skills, and training, (5) barriers to housing and services, (6) crime and disorder, and (7) living environment. LSOAs comprise between 400 and 1200 households and typically contain a resident population between 1000 and 3000 persons. LSOA boundaries remain fixed over time, ensuring that values of the IMD are comparable over time. National quintiles of area deprivation are created through ranking LSOAs according to their deprivation score. The postcode address of responding households in each survey was linked to the LSOA, which was then used to determine the corresponding deprivation quintile. The IMD was first included in the HSE 2004 dataset and was updated in 2007, 2010, and 2015; the HSE datasets available at the UK Data Service (and the harmonized dataset compiled for our study) contain the version of the IMD that was current at the time of each survey.

### *Behavioral Risk Factors: PA and Alcohol*

In the HSE series, questions on PA assessed frequency (number of days spent doing a specified activity in the last 4 weeks) and duration (of an average episode lasting above a specified bout duration limit) in 4 leisure-time domains: domestic activity, do-it-yourself or manual work, walking, and sports or exercise. In the reporting of trends, PA undertaken while at work is also considered in the estimation of summary activity levels for HSE reports. PAs are classified into intensity levels (light, moderate, and vigorous) based on an estimate of the energy expenditure associated with each activity.

Changes in the PA questions (reflecting changes over time in policy recommendations, namely, the reference period for bouts of activities to report) have restricted the meaningfulness of comparisons over time to some extent. The lower duration limit for an activity to be included was 15 minutes in 1998 and 2006; 30 minutes in 2003 (15 minutes for sports and exercise); and 10 minutes in 2008, 2012, and 2016. A single question on

occupational PA ("Thinking about your job, in general would you say that you are very physically active, fairly physically active, not very physically active, or not at all physically active?") was asked in 2003 and 2006; more detailed questions introduced in 2008 (repeated in 2012 and 2016) focused on what people actually do at work (eg, climbing stairs or ladders, lifting, and carrying or moving heavy loads) and how many hours they typically work.

To maximize the trend series, we derived a variable summarizing the number of days per week that participants undertook PA of at least moderate intensity for a minimum duration of 30 minutes. For those participants who reported that they were very or fairly active in their job, arbitrary estimates of 12 or 20 working days in the last 4 weeks (3 or 5 days per week, respectively) were used, depending on whether the participant worked part time or full time, to assess levels of PA while at work.

The main interview included questions on the number of drinking days in the last week (collected in all years), alcohol consumption (type and quantity) on the heaviest drinking day in the last week (all years), and average weekly drinking over the past 12 months (2011 onward). Information on the type and quantity of drinks consumed were used to estimate alcohol unit consumption using a method of conversion detailed elsewhere [57]. The applied conversion factors were revised in 2006 to 2007 to account for changes to the drinking environment. Alcohol units were categorized to represent consumption on the heaviest drinking day relative to recommended daily limits at the time of the survey (>3 units for women and >4 units for men); binge drinking was defined as drinking twice the recommended daily limits (>6 units for women and >8 units for men) [58]. Additional variables classified participants according to whether they drink alcohol nowadays (2 categories: nondrinker and current drinker; 3 categories: never, former, and current drinker).

### General Health and Long-Standing Illness

Participants were asked to rate their health in general (response options: very good, good, fair, bad, and very bad). Long-standing illnesses were also reported in the survey. Before 2012, the question on long-standing illness referred to "an illness, disability or infirmity...that has troubled you over a period of time or that is likely to affect you over a period of time." Since 2012, long-standing illness is defined as "any physical or mental health condition or illness lasting or expected to last 12 months or more."

### Diagnosed CVD Conditions

The HSE surveys for 1998, 2003, 2006, 2011, and 2017 had a specific focus on CVD. During the interview, adults were asked a series of questions about whether they had ever been diagnosed with certain specified CVDs, and if so, whether the diagnosis had been made by a physician. The specified conditions included angina, myocardial infarction, stroke, abnormal heart rhythm, a heart murmur or "other cardiovascular condition." No attempt was made to verify these self-reported diagnoses. Therefore, it is possible that some misclassification may have occurred because some participants may not have remembered, or may have misremembered, the diagnosis made by their physician.

### Use of Medicines

At the nurse visit, participants were asked the following: "Are you taking or using any medicines, pills, syrups, ointments, puffers or injections prescribed for you by a doctor or nurse?" Those who did were then asked the name of each prescribed item. In most cases, participants showed the nurse the actual medicine pack. These were coded by the nurse into medicine classes based on the subsections of the British National Formulary. Up to 22 medicines could be recorded (this has recently increased to 32). For each medicine, a follow-up question asked whether they had taken or used that medicine in the last 7 days. Variables on the use of CVD medicines, lipid-lowering medicines, and BP-lowering medicines are provided in the harmonized dataset.

### Pregnancy Status

At the nurse visit, women aged 16 to 49 years were asked whether they were pregnant at the moment.

### Contraceptive Use

Some questions were completed by the participants in paper self-completion questionnaires. In the HSE 1998, 2001 to 2003, and 2005 to 2006, this included questions for women on whether they had ever taken the contraceptive pill or had a contraceptive injection or implant. Those replying yes were asked whether they were currently taking the contraceptive pill or having a contraceptive injection or implant. On the basis of these 2 questions, we created a three-category variable distinguishing between women who reported that they (1) had never taken the contraceptive pill or had a contraceptive injection or implant, (2) had ever taken but were not currently taking the contraceptive pill or having a contraceptive injection or implant, and (3) those currently taking the contraceptive pill or having a contraceptive injection or implant. In addition, the current use of oral contraceptives was recorded each year at the nurse visit in the use of medicines section.

### Other Variables

Other sociodemographic variables compiled in the harmonized dataset included marital status (single, married, separated, divorced, widowed, and cohabitees), ethnic group (White, Black, Asian, mixed, and other), government office region (GOR: North East, North West, Yorkshire and the Humber, East Midlands, West Midlands, East of England, London, South East, and South West), an urban or rural indicator, and receipt of various means-tested state benefits (eg, Income Support and Housing Benefit).

### Sampling Design Information (Primary Sampling Units, Strata, and Weights)

Using the small-user Postcode Address File as the sampling frame, a 2-stage stratified random sampling process was used to select each year's general population sample. First, a random sample of primary sampling units (PSUs), based on postcode sectors, was selected, with probability proportional to the total number of addresses. Stratification was performed by ordering the PSUs according to local authority, and within each local authority by the percentage of households in the last census where the head of household was in a nonmanual occupation. The list of PSUs was then sampled at fixed intervals from a

random starting point. Second, a random sample of a fixed number of addresses was then drawn from each PSU, ensuring a self-weighted design in which every eligible participant had the same probability of selection.

Each pair of PSUs in the ordered list was assigned to the same stratum. Since 2006, the Taylor series method (linearization) has been used in annual HSE reporting for variance estimation using the PSU and stratum identifiers. For the analyses of data pooled over several years, GOR has often been used as an alternative stratification variable.

In 2003, weighting the general population adult sample for nonresponse was introduced for the first time in the HSE series [59]. The nonresponse weights take account of nonresponse at 4 levels: household response, individual response to the interview, individual response to the nurse visit, and individual response to the collection of blood samples. The harmonized dataset includes the relevant interview, nurse, and blood sample weights for each survey year from 2003 onward. These weights are scaled so that their sum over the relevant set of participants equals the unweighted sample size (resulting in an average weight of 1); the weighting variables before 2003 were assigned the value 1.

## Results

### Analytical Samples

A total of 190,905 adults (aged ≥16 years) from the general population samples completed the health interview between 1998 and 2017 (Figures 1 and 2). The harmonized dataset excludes the participants in the boost years of HSE 1999, 2000, and 2004 (22,490/190,905, 11.78%) but includes the boost sample of adults aged ≥65 years in HSE 2005 (2673/193,578, 1.38%), resulting in a provided dataset of 88.38% (171,088/193,578) adults. Excluding the boost sample of adults aged ≥65 years in HSE 2005 for this study produced a dataset of 168,415 (nonboost sample) adults, of which 75,980 (45.12%) were excluded from the analyses due to falling outside the age range of 40 to 74 years.

**Figure 1.** Flowchart of participants included in the estimation of changes over time in cardiovascular disease (CVD) risk (laboratory-based scores).*Allocated to physical activity module; **allocated to CVD (including diabetes) module.

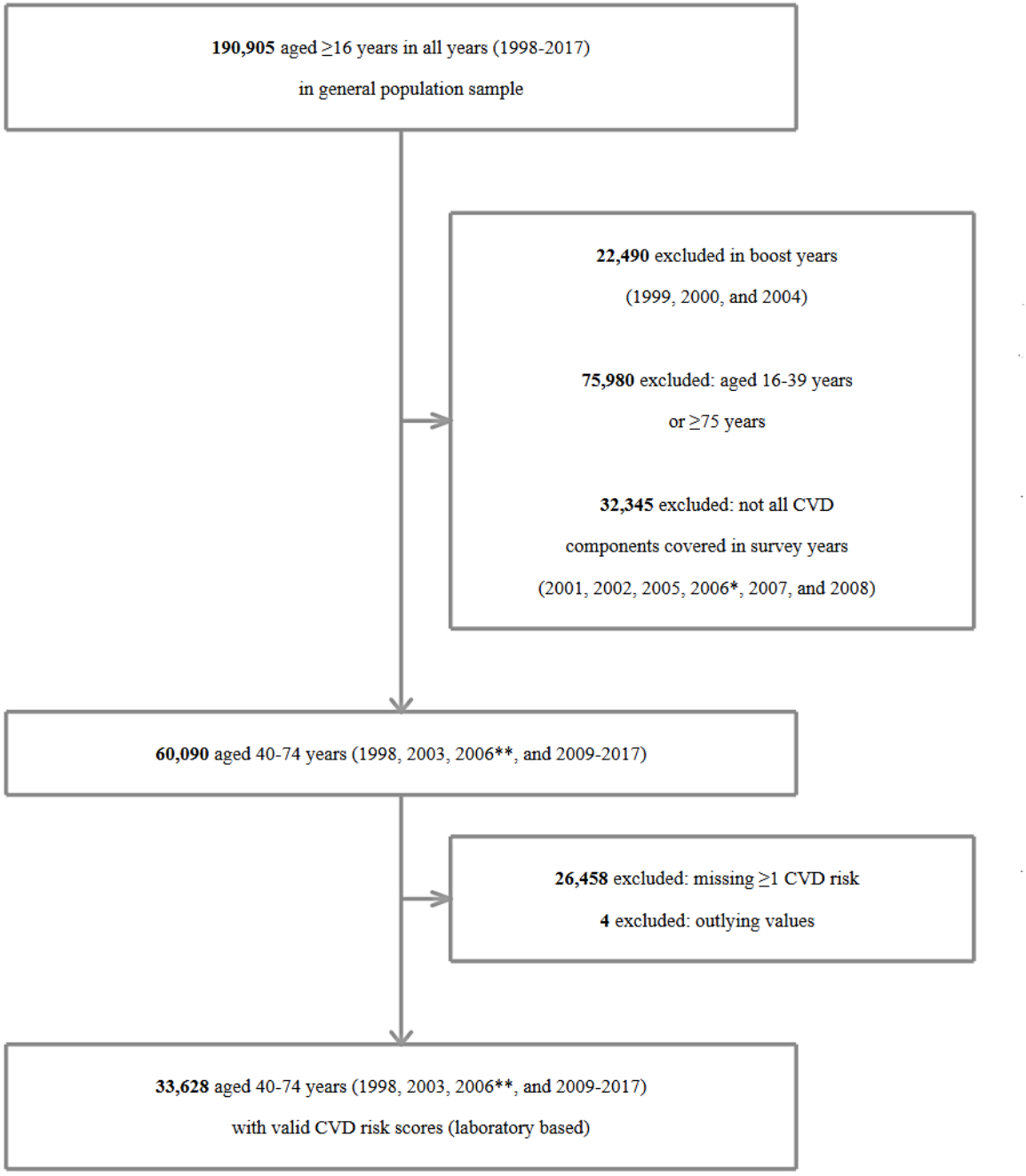
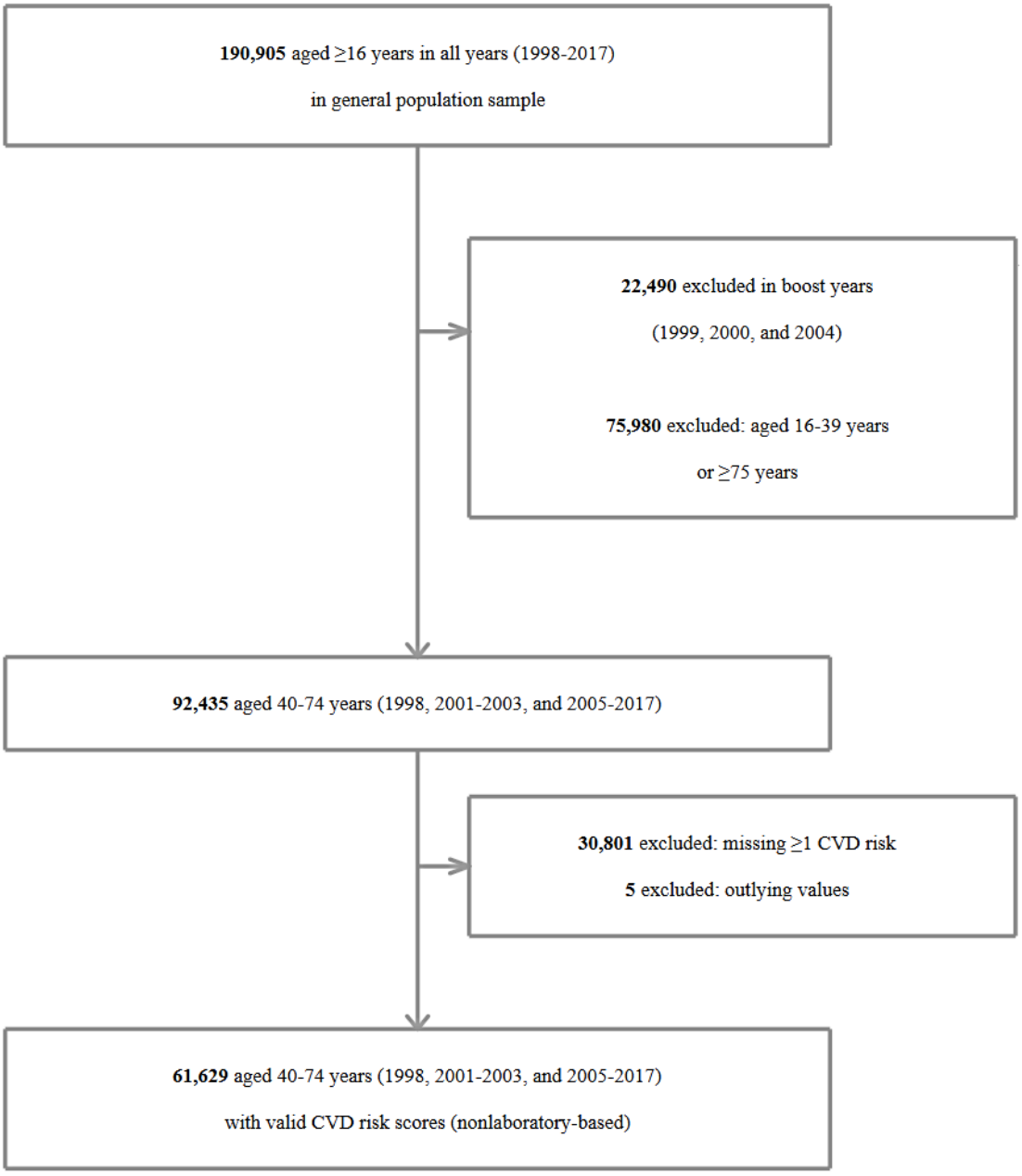


**Figure 2.** Flowchart of participants included in the estimation of changes over time in cardiovascular disease (CVD) risk (non–laboratory-based scores).




## Missing Data on CVD Risk Scores

As shown in Figures 1 and 2, in the years when all CVD risk components were included in the survey, a sizeable number of adults aged 40 to 74 years were excluded from the analyses due to missing data on at least 1 risk component (30,801/92,435, 33.32% and 26,458/60,090, 44.03% for the non–laboratory-based and laboratory-based risk scores, respectively). The calculation of CVD risk scores requires complete (ie, nonmissing) risk factor information. As SBP is a component of both algorithms, inclusion in the analytical samples for calculating CVD risk is contingent on participants having participated in the nurse-visit stage of the survey and having their BP measured. In addition, as total cholesterol is a component of the laboratory-based scores, inclusion in this analytical sample is contingent on participants providing a nonfasting blood sample. Nonparticipation in the nurse visit and blood sample collection is therefore the main driver for the amount of missing data shown in the final stage of the flowcharts provided in Figures 1 and 2. An additional factor contributing to missing data for the non–laboratory-based scores is missing BMI data, due to refusals to undergo weight measurement during the health interview.

For the participants with complete and valid (ie, nonoutlying) data on each individual risk component, laboratory-based and non–laboratory-based 10-year CVD risk scores were calculated (33,628/60,090, 55.96% and 61,629/92,435, 66.67% participants aged 40 to 74 years, respectively). On the basis of unweighted data, the mean age of participants with laboratory-based scores was 56.1 (SD 9.8) years; 54.11% (18,197/33,628) of the participants were female. The sociodemographic profile was similar for those with non–laboratory-based scores.

## Analysis Plan

Analyses were performed separately by sex, given notable differences in CVD risk. These were conducted using Stata (version 18.0; StataCorp) with survey analysis procedures to account for the complex survey design (PSUs; GOR [strata]; and appropriate nonresponse weights, ie, nurse weights for the non–laboratory-based sample and blood sample weights for the laboratory-based sample).

For each survey year, we estimated the percentages (diagnosed diabetes and current smoking) and means of the individual risk components and the mean predicted 10-year risk of CVD (Figures 3 and 4). Wald tests were performed to test the null hypothesis of no change in the mean predicted 10-year risk of CVD between the first and last survey periods (1998 and 2017, respectively). Linear trends in CVD risk were tested using linear regression, with the predicted risk score as the outcome and survey year (continuous variable) as the independent variable. Statistical tests were 2-sided, and $P<.05$ was considered statistically significant.

**Figure 3.** A 10-year cardiovascular disease (CVD) risk score (laboratory based) and its components by survey year and sex. SBP: systolic blood pressure; TC: total cholesterol; WHO: World Health Organization.

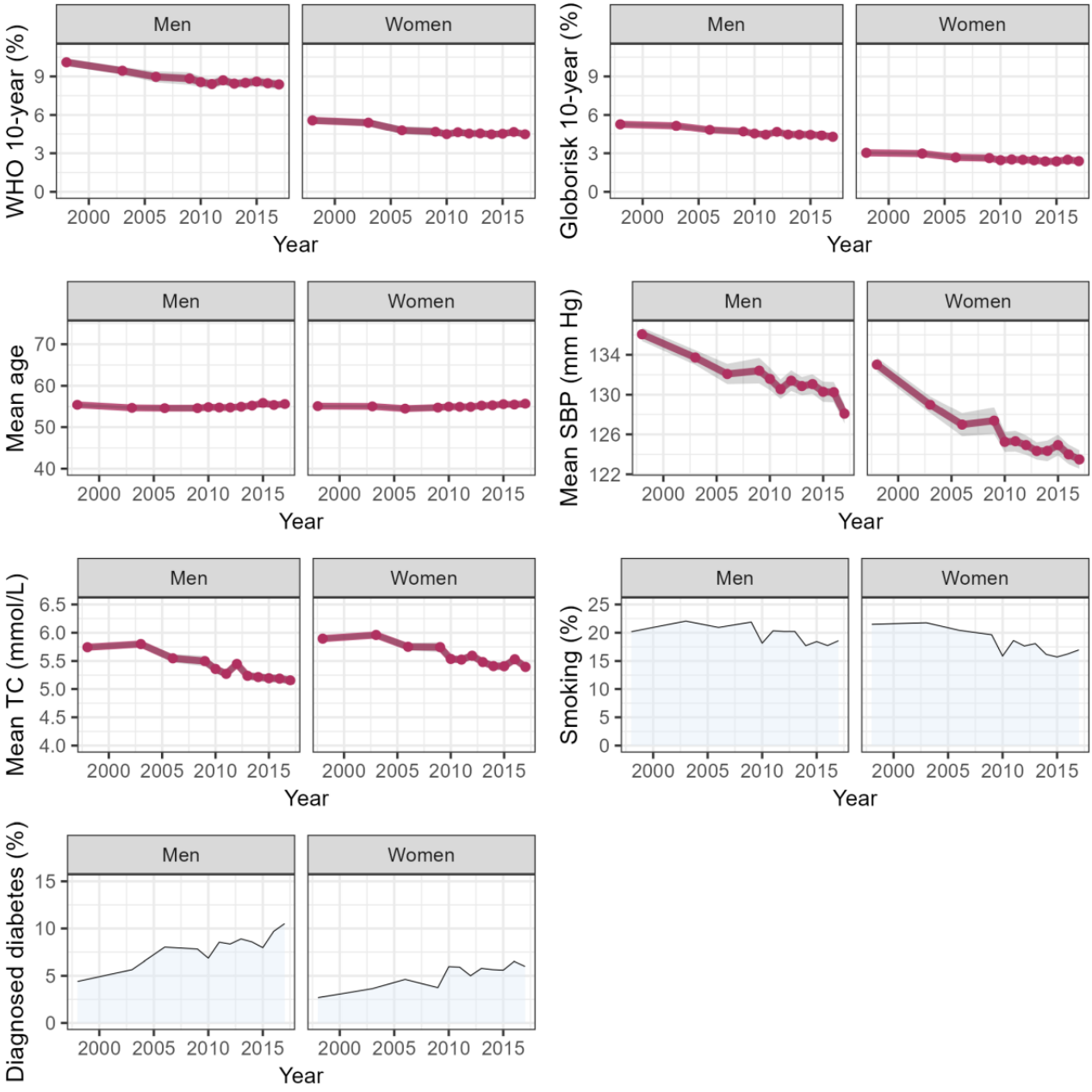
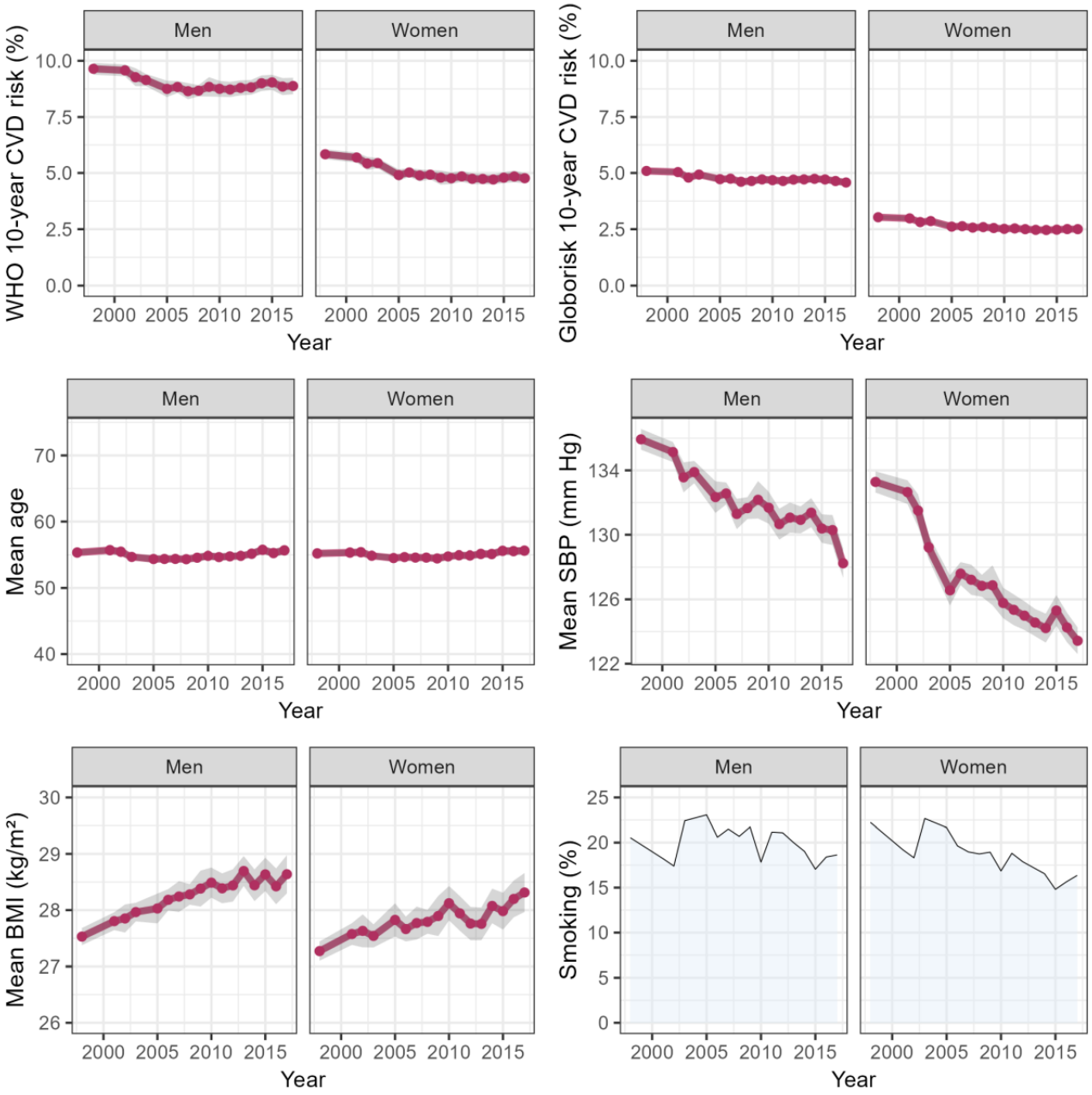


**Figure 4.** A 10-year cardiovascular disease (CVD) risk score (non–laboratory-based) and its components by survey year and sex. SBP: systolic blood pressure; WHO: World Health Organization.




## Trends in CVD Risk

The mean predicted 10-year CVD risk declined significantly over the last 2 decades in both sexes (for Wald tests, all $P \leq .001$; for linear trend, all $P < .001$; Table 2). In men, the mean of the laboratory-based WHO risk score was 10.1% (SE 0.2%) and 8.4% (SE 0.2%) in 1998 and 2017, respectively; corresponding figures in women were 5.6% (SE 0.1%) and 4.5% (SE 0.1%). In men, the mean of the non–laboratory-based WHO risk score was 9.6% (SE 0.1%) and 8.9% (SE 0.2%) in 1998 and 2017, respectively; corresponding figures in women were 5.8% (SE 0.1%) and 4.8% (SE 0.1%). Globorisk risk scores were lower in absolute terms, but the pattern of change was very similar (for linear trend, all $P < .001$).

**Table 2.** Estimated linear trend in 10-year cardiovascular disease risk, Health Survey for England data (1998-2017).

| | WHO[a] | | Globorisk | |
|---|---|---|---|---|
| | β (%[b]; 95% CI) | *P* value[c] | β (%; 95% CI) | *P* value |
| **Laboratory based** | | | | |
| Men | –0.09 (–0.11 to –0.07) | <.001 | –0.05 (–0.06 to –0.05) | <.001 |
| Women | –0.06 (–0.08 to –0.05) | <.001 | –0.04 (–0.04 to –0.03) | <.001 |
| **Non–laboratory based** | | | | |
| Men | –0.04 (–0.06 to –0.03) | <.001 | –0.02 (–0.03 to –0.02) | <.001 |
| Women | –0.06 (–0.07 to –0.05) | <.001 | –0.03 (–0.04 to –0.03) | <.001 |

[a]WHO: World Health Organization.

[b]Linear trends in CVD risk were tested using linear regression (accounting for the complex survey design), with the risk score as the outcome and survey year (continuous variable) as the predictor. The slope (β coefficient) represents the estimated annual decrease in the mean 10-year CVD risk (in absolute terms, expressed as a percentage). For example, for the laboratory-based WHO algorithm, the estimated annual decrease in the predicted 10-year CVD risk for men was 0.09% (eg, from 9.94% in 1998 to 9.85% in 1999).

[c]*P* value for linear trend.

### Trends in CVD Risk Components

The significantly declining linear trends in the mean predicted 10-year CVD risk reflected the net effect of diverging trends in its risk components. On the one hand, the data showed significant declines between the first and last survey periods in mean SBP (2017 vs 1998: declines of 8 mm Hg and 10 mm Hg in men and women, respectively), mean total cholesterol (0.6 mmol/L and 0.5 mmol/L), and lower levels of current smoking (decrease of 5 percentage points [PPs] in women; for Wald tests, all $P \leq .001$; except $P=.002$ for smoking in women). Simultaneously, significant increases occurred in mean BMI (2017 vs 1998: increases of 1.1 kg/m$^2$ and 1.0 kg/m$^2$ in men and women, respectively) and levels of diagnosed diabetes (6 PPs and 3 PPs in men and women, respectively; for Wald tests, all $P \leq .001$).

## Discussion

### Principal Findings

As CVDs remain the leading cause of death globally, using nationally representative health surveys from a high-income country such as England to model temporal trends in CVD risk can provide guidance for middle-income countries such as South Africa to inform where best to intervene and direct resources to reduce disease burden.

Modeling temporal trends in CVD risk requires pooling annual cross-sectional health surveys. Compiling and appending data from repeated cross-sectional surveys to enable such modeling is a daunting task due to changes in aspects such as survey content, question wording, inclusion of boost samples, weighting, measuring equipment, and guidelines for data protection. While data harmonization across aging cohorts such as the US Health and Retirement Study and the English Longitudinal Study of Ageing has benefitted enormously from the efforts of the Gateway to Global Aging team (including the production of harmonized datasets) [60], no such platform exists to enable researchers to harmonize data across repeated cross-sections of health examination surveys such as the HSE. In this manuscript, we have documented the methods and procedures used to painstakingly compile the harmonized dataset based on 17 years of separate HSE datasets spanning 2 decades (1998-2017), including a description of how we calculated the predicted 10-year risk of CVD using the WHO [35] and Globorisk [36-38] CVD risk algorithms.

In our presentation of early results, we showed significant declines over time in the mean predicted 10-year total (ie, fatal and nonfatal) CVD risk in both sexes, suggesting an improvement in cardiovascular health at the population level, consistent with modeling studies in England pointing to the role of increased prevention and treatment [61,62]. The observed trends in CVD risk reflect the net effect of divergent trends in its risk components, namely, significant declines in average levels of SBP, total cholesterol, and current smoking (women only), with simultaneous increases in mean BMI and diagnosed diabetes. This complex pattern of temporal trends in the individual CVD risk components agrees with other studies using HSE data over the same period [63].

### Implications of Our Findings

In the later stages of the EXPOSE study, more complex regression techniques will be used to compare trends in CVD risk between South Africa and England and empirically test the relative contributions of a wide set of factors that may explain those trends, including demographic, behavioral, social, environmental, and health care–related aspects. How the findings of this study apply to different countries is likely to be influenced by socioeconomic structures and health care systems (eg, access to health care is free at the point of use in the United Kingdom). Bearing this caveat in mind, our initial findings on the significant declines in 10-year CVD risk over 2 decades, accompanied by the conflicting trends in its modifiable risk components, can be leveraged to inform public health policy and interventions in the United Kingdom and in low- and middle-income countries such as South Africa with high CVD burdens.

First, our descriptive analyses show that the significant declines in the predicted 10-year risk for CVD may be attributable in

large measure to population-level declines in cigarette smoking and in mean levels of BP and total cholesterol. In the absence of increasing levels of diagnosed diabetes and BMI, predicted risk would have declined at a stronger pace.

Second, the favorable trends in CVD risk demonstrates the population-level gains in cardiovascular health that are achievable through implementing a wide range of population-based public health primary and secondary prevention approaches. These include (1) policy and regulatory measures (eg, tobacco taxation and antismoking legislation, including smoke-free workplaces and public places); (2) public health campaigns promoting awareness about lifestyle behaviors (eg, diet and exercise); and (3) improvements in the early detection and management of CVD-related conditions such as hypertension, dyslipidemia, and diabetes through initiatives such as the National Health Service Health Check program and financial incentivization of general practices in screening for individual CVD risk factors (eg, increasing use of antihypertensive medicines and statins). Building on these successes, low- and middle-income countries could adopt similar approaches, adjusting for local socioeconomic and cultural contexts.

Third, evidence on the increasing levels of diagnosed diabetes and BMI shows that substantial challenges remain in reducing the CVD burden, and this can be used to leverage the expansion of prevention efforts to include combined lifestyle interventions to improve diets, levels of PA, and achieve sustained weight loss.

Finally, our study demonstrates the availability of long-standing, high-quality, nationally representative health examination survey data in high-income countries such as England to monitor population trends in CVD risk and its components, offering valuable evidence to inform public health policy, guide resource allocation, design targeted prevention strategies, and assess their effectiveness. Building similar capacity in population health surveillance in low- and middle-income countries is a major challenge due to factors such as budgetary constraints [64], but such investment would greatly contribute to identifying priorities for CVD prevention and evaluating the success of interventions.

### Strengths and Limitations

Our study uses high-quality data on the individual components of CVD risk, including objective measurements of BP, total cholesterol, and BMI, which avoids the potential inaccuracies of self-reported measures. Participants from health examination surveys such as the HSE are not selected on the basis of health care use, thereby increasing representativeness and avoiding selection bias to some extent. The harmonized dataset covers a time span of 2 decades, enabling modeling of temporal trends in CVD risk and investigation of which factors explain the trends. Area-level variables such as relative deprivation and urbanicity are also provided with the dataset, permitting analysis of contextual effects.

Although the authors of this study have considerable experience in collecting and analyzing HSE data, creating a harmonized dataset was a daunting task. The accuracy of variable derivation (eg, appropriate recoding to ensure congruence of the values across datasets) was checked by comparing estimates with the available trend tables published in annual HSE reports. We hope that the dissemination of our methods and procedures as well as the provision of code for harmonizing and appending the annual datasets will support future efforts by the wider research community.

Limitations of our study include increasing levels of nonresponse and reliance on complete case analyses in our presentation of early results (possibly biasing results). As mentioned earlier, the calculation of CVD risk scores requires complete (ie, nonmissing) risk factor information, and this approach is consistent with the model derivation stage of algorithms such as the WHO and Globorisk, which excluded participants with missing data on any of the selected risk factors.

As age in single-year intervals is no longer provided on the EUL datasets (to preserve the anonymity of participants), the calculation of predicted CVD risk using the midpoint of categorical age (in 5-year intervals) for participants in HSE 2016 to 2017 has inevitably reduced precision to some extent. A final limitation of our study is the cross-sectional nature of the HSE design, which prevents any validation of the risk algorithms (in the absence of appropriate data linkages).

### Conclusions

Monitoring temporal trends in predicted CVD risk and its risk factors at the population level is vital to support prevention efforts. Alongside evidence from longitudinal databases, harmonized data from repeated cross-sectional nationally representative health surveys can be used to identify and quantify the drivers of recent changes in CVD risk.

## Acknowledgments

The authors would like to thank the interviewers and nurses, the participants in the Health Survey for England series, and colleagues at NatCen Social Research. The authors would also like to thank the National Health Service England. The EXPOSE (Explaining Population Trends in Cardiovascular Risk: A Comparative Analysis of Health Transitions in South Africa and England) study was supported by a scientific grant from the Economic and Social Research Council (ES/V003259/1; principal investigator: KA-G).

## Authors' Contributions

SS contributed to the conceptualization, methodology, software, validation, formal analysis, data curation, writing the original draft, and visualization. JSM participated in validation, reviewing and editing the draft, and project administration. MT-S was

involved in validation and reviewing and editing the draft. AC contributed to the conceptualization, methodology, and reviewing and editing the draft. KA-G played a role in conceptualization, methodology, software, validation, data curation, reviewing and editing the draft, supervision, project administration, and funding acquisition.

## Conflicts of Interest

None declared.

## References


1. Global status report on noncommunicable diseases 2010. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/44579/9789240686458_eng.pdf [accessed 2024-10-24]
2. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1736-1788 [FREE Full text] [doi: 10.1016/S0140-6736(18)32203-7] [Medline: 30496103]
3. Cardiovascular diseases (CVDs). World Health Organization. 2021 Jun 11. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2024-10-24]
4. Adjaye-Gbewonyo K, Cois A. Explaining population trends in cardiovascular risk: protocol for a comparative analysis of health transitions in South Africa and England using nationally representative survey data. BMJ Open 2022 Mar 29;12(3):e061034 [FREE Full text] [doi: 10.1136/bmjopen-2022-061034] [Medline: 35351734]
5. ExPoSE homepage. ExPoSE. URL: https://www.exposeproject.net/ [accessed 2024-10-24]
6. Mindell J, Biddulph JP, Hirani V, Stamatakis E, Craig R, Nunn S, et al. Cohort profile: the Health Survey for England. Int J Epidemiol 2012 Dec;41(6):1585-1593. [doi: 10.1093/ije/dyr199] [Medline: 22253315]
7. Explaining population trends in cardiovascular risk in South Africa and England 1998-2017. DataFirst. 2024 Sep 11. URL: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/981 [accessed 2024-12-26]
8. Craig R, Mindell J, Vasant H. Health Survey for England 2008. Volume 1: physical activity and fitness. The NHS Information Centre for Health and Social Care. 2008. URL: https://tinyurl.com/4fzjefrp [accessed 2024-12-20]
9. Health Survey for England - 2010, respiratory health. National Health Service England. 2011 Dec 15. URL: https://tinyurl.com/7bzsrdh3 [accessed 2024-12-17]
10. Erens B, Primatesta P, Prior G. Health Survey for England 1999: The Health of Minority Ethnic Groups. London, UK: The Stationery Office; 2001.
11. Sproston K, Mindell J. Health Survey for England 2004. Volume 1: The health of minority ethnic groups. The Information Centre. 2006. URL: https://tinyurl.com/5vms6ee5 [accessed 2024-12-20]
12. Health Survey for England 2000: The Health of Older People. London, UK: The Stationery Office; 2002.
13. Sproston K, Primatesta P. Health Survey for England 2002. Volume 1: The Health of Children and Young People. London, UK: The Stationery Office; 2003.
14. Health Survey for England 2005, Health of Older People. National Health Service England. 2007 Mar 23. URL: https://tinyurl.com/ytjjrahv [accessed 2024-12-20]
15. Margozzini P, Tolonen H, Bernabe-Ortiz A, Cuschieri S, Donfrancesco C, Palmieri L, et al. National health examination surveys: an essential piece of the health planning puzzle. medRxiv Preprint posted online on July 13, 2023 [FREE Full text] [doi: 10.1101/2023.07.11.23292221] [Medline: 37503238]
16. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 1998. [data collection]. 5th release. UK Data Service. 2010. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=4150#!#1 [accessed 2024-12-16]
17. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2000. [data collection]. 4th release. UK Data Service. 2011. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=4487#!#1 [accessed 2024-12-16]
18. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2001. [data collection]. 3rd release. UK Data Service. 2010. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=4628#!#1 [accessed 2024-12-16]
19. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2002. [data collection]. 2nd release. UK Data Service. 2010. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=4912#!#1 [accessed 2024-12-16]
20. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2003. [data collection]. 2nd release. UK Data Service. 2010. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=5098#!#1 [accessed 2024-12-16]
21. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2005. [data collection]. 3rd release. UK Data Service. 2011. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=5675#!#1 [accessed 2024-12-16]

22. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2006. [data collection]. 4th release. UK Data Service. 2011. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=5809#!#1 [accessed 2024-12-16]
23. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2007. [data collection]. 2nd release. UK Data Service. 2010. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6112#!#1 [accessed 2024-12-16]
24. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2008. [data collection]. 4th release. UK Data Service. 2013. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6397#!#2 [accessed 2024-12-16]
25. National Centre for Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2009. [data collection]. 3rd release. UK Data Service. 2015. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6732#!#2 [accessed 2024-12-16]
26. NatCen Social Research, Department of Epidemiology and Public Health, Royal Free and University College Medical School. Health survey for England, 2010. [data collection]. 3rd release. UK Data Service. 2015. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6986#!#3 [accessed 2024-12-16]
27. NatCen Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2011. [data collection]. 2nd edition. UK Data Service. 2023. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=7260#!#1 [accessed 2024-12-16]
28. NatCen Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2012. [data collection]. 2nd edition. UK Data Service. 2023. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=7480#!#1 [accessed 2024-12-16]
29. NatCen Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2013. [data collection]. 2nd edition. UK Data Service. 2023. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=7649#!#1 [accessed 2024-12-16]
30. NatCen Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2014. [data collection]. 4th edition. UK Data Service. 2023. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=7919#!#3 [accessed 2024-12-16]
31. NatCen Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2015. [data collection]. 3rd edition. UK Data Service. 2023. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=8280#!#1 [accessed 2024-12-16]
32. NatCen Social Research, Department of Epidemiology and Public Health, University College London. Health survey for England, 2016. [data collection]. 4th edition. UK Data Service. 2023. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=8334#!#1 [accessed 2024-12-16]
33. Department of Epidemiology and Public Health, University College London, National Centre for Social Research. Health survey for England, 2017. [data collection]. 3rd edition. UK Data Service. 2023. URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=8488#!#3 [accessed 2024-12-16]
34. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J 1991 Jan;121(1 Pt 2):293-298 [FREE Full text] [doi: 10.1016/0002-8703(91)90861-b] [Medline: 1985385]
35. WHO CVD Risk Chart Working Group. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health 2019 Oct;7(10):e1332-e1345 [FREE Full text] [doi: 10.1016/S2214-109X(19)30318-3] [Medline: 31488387]
36. Hajifathalian K, Ueda P, Lu Y, Woodward M, Ahmadvand A, Aguilar-Salinas CA, et al. A novel risk score to predict cardiovascular disease risk in national populations (Globorisk): a pooled analysis of prospective cohorts and health examination surveys. Lancet Diabetes Endocrinol 2015 May;3(5):339-355. [doi: 10.1016/s2213-8587(15)00081-9]
37. Ueda P, Woodward M, Lu Y, Hajifathalian K, Al-Wotayan R, Aguilar-Salinas CA, et al. Laboratory-based and office-based risk scores and charts to predict 10-year risk of cardiovascular disease in 182 countries: a pooled analysis of prospective cohorts and health surveys. Lancet Diabetes Endocrinol 2017 Mar;5(3):196-213. [doi: 10.1016/s2213-8587(17)30015-3]
38. Cohorts Consortium of Latin Americathe Caribbean (CC-LAC). Derivation, internal validation, and recalibration of a cardiovascular risk score for Latin America and the Caribbean (Globorisk-LAC): a pooled analysis of cohort studies. Lancet Reg Health Am 2022 May;9:None [FREE Full text] [doi: 10.1016/j.lana.2022.100258] [Medline: 35711683]
39. Farzadfar F, Finucane MM, Danaei G, Pelizzari PM, Cowan MJ, Paciorek CJ, et al. National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3·0 million participants. The Lancet 2011 Feb;377(9765):578-586. [doi: 10.1016/s0140-6736(10)62038-7]
40. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet 2017 Dec 16;390(10113):2627-2642 [FREE Full text] [doi: 10.1016/S0140-6736(17)32129-3] [Medline: 29029897]

41. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19·1 million participants. Lancet 2017 Jan 07;389(10064):37-55 [FREE Full text] [doi: 10.1016/S0140-6736(16)31919-5] [Medline: 27863813]
42. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet 2016 Apr 09;387(10027):1513-1530 [FREE Full text] [doi: 10.1016/S0140-6736(16)00618-8] [Medline: 27061677]
43. GBD 2016 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1345-1422 [FREE Full text] [doi: 10.1016/S0140-6736(17)32366-8] [Medline: 28919119]
44. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
45. Murray CJ, Ezzati M, Flaxman AD, Lim S, Lozano R, Michaud C, et al. GBD 2010: design, definitions, and metrics. The Lancet 2012 Dec;380(9859):2063-2066. [doi: 10.1016/s0140-6736(12)61899-6]
46. Boyer C, Danaei G, Hajifathalian K, Ueda PM, Rodrigo MC. globorisk: global CVD risk calculator. GitHub. 2022. URL: https://christopherbboyer.com/globorisk/ [accessed 2024-10-25]
47. Scholes S, Fat LN, Mindell JS. Trends in cardiovascular disease risk factors by BMI category among adults in England, 2003-2018. Obesity (Silver Spring) 2021 Aug 21;29(8):1347-1362. [doi: 10.1002/oby.23184] [Medline: 34155827]
48. BMI classification. World Health Organization. 2006. URL: http://www.assessmentpsychology.com/icbmi.htm [accessed 2024-10-26]
49. Falaschetti E, Mindell J, Knott C, Poulter N. Hypertension management in England: a serial cross-sectional study from 1994 to 2011. The Lancet 2014 May;383(9932):1912-1919. [doi: 10.1016/s0140-6736(14)60688-7]
50. Erens B, Bajekal M, Primatesta P. Health Survey for England: Cardiovascular Disease '98. London, UK: The Stationery Office; 1999.
51. Scholes S, Conolly A, Mindell JS. Income-based inequalities in hypertension and in undiagnosed hypertension: analysis of Health Survey for England data. J Hypertens 2020 May;38(5):912-924. [doi: 10.1097/HJH.0000000000002350] [Medline: 31913219]
52. Moody A, Cowley G, Ng Fat L, Mindell JS. Social inequalities in prevalence of diagnosed and undiagnosed diabetes and impaired glucose regulation in participants in the Health Surveys for England series. BMJ Open 2016 Feb 08;6(2):e010155 [FREE Full text] [doi: 10.1136/bmjopen-2015-010155] [Medline: 26857106]
53. Health Survey for England 2019 [NS]. National Health Service England. 2020 Dec 15. URL: https://tinyurl.com/33jzmkds [accessed 2024-12-20]
54. HbA1c conversion table. Australian Diabetes Society. URL: https://tinyurl.com/mtrj5b3m [accessed 2024-10-26]
55. Health Survey for England 2015: user guide. NatCen Social Research. URL: https://tinyurl.com/nxvtb4az [accessed 2024-12-17]
56. McClements LD. Equivalence scales for children. J Public Econ 1977 Oct;8(2):191-210. [doi: 10.1016/0047-2727(77)90018-4] [Medline: 12278685]
57. Fuller E. Alcohol consumption. In: Health Survey for England 2006: Volume 1: Cardiovascular Disease and Risk Factors in Adults. London, UK: The Information Centre; 2008.
58. Knott CS, Scholes S, Shelton NJ. Could more than three million older people in England be at risk of alcohol-related harm? A cross-sectional analysis of proposed age-specific drinking limits. Age Ageing 2013 Sep;42(5):598-603. [doi: 10.1093/ageing/aft039] [Medline: 23880141]
59. Sproston K, Primatesta P. Health Survey for England 2003: Volume 2: Risk Factors for Cardiovascular Disease. London, UK: The Stationery Office; 2004.
60. Gateway to global aging data homepage. Gateway to Global Aging Data. URL: https://g2aging.org/home [accessed 2024-08-01]
61. Bajekal M, Scholes S, Love H, Hawkins N, O'Flaherty M, Raine R, et al. Analysing recent socioeconomic trends in coronary heart disease mortality in England, 2000-2007: a population modelling study. PLoS Med 2012 Jun 12;9(6):e1001237 [FREE Full text] [doi: 10.1371/journal.pmed.1001237] [Medline: 22719232]
62. Guzman-Castillo M, Ahmed R, Hawkins N, Scholes S, Wilkinson E, Lucy J, et al. The contribution of primary prevention medication and dietary change in coronary mortality reduction in England between 2000 and 2007: a modelling study. BMJ Open 2015 Jan 22;5(1):e006070 [FREE Full text] [doi: 10.1136/bmjopen-2014-006070] [Medline: 25613952]
63. Montano D. Education differences in cardiometabolic risk in England, Scotland and the United States between 1992 and 2019. BMC Cardiovasc Disord 2022 Jun 02;22(1):247 [FREE Full text] [doi: 10.1186/s12872-022-02681-y] [Medline: 35655138]

64. Echouffo-Tcheugui JB, Yaya S, Joshi R, Narayan KM, Kengne AP. Population surveillance of cardiovascular diseases in low-income to middle-income countries should leverage existing international collaborations. BMJ Glob Health 2018 Oct 01;3(5):e000866 [FREE Full text] [doi: 10.1136/bmjgh-2018-000866] [Medline: 30294457]

## Abbreviations

**BP:** blood pressure
**CHD:** coronary heart disease
**CVD:** cardiovascular disease
**EUL:** end-user license
**EXPOSE:** Explaining Population Trends in Cardiovascular Risk: A Comparative Analysis of Health Transitions in South Africa and England
**GBD:** Global Burden of Disease
**GOR:** government office region
**HbA1c:** glycated hemoglobin
**HR:** hazard ratio
**HSE:** Health Survey for England
**IMD:** index of multiple deprivation
**LSOA:** lower-layer super output area
**PA:** physical activity
**PP:** percentage point
**PSU:** primary sampling unit
**SBP:** systolic blood pressure
**SES:** socioeconomic status
**SL:** special license
**WHO:** World Health Organization

*Edited by A Coristine; submitted 01.08.24; peer-reviewed by L Ng Fat, WF Khaw; comments to author 29.09.24; revised version received 02.11.24; accepted 09.12.24; published 20.01.25.*

*Please cite as:*
*Scholes S, Mindell JS, Toomse-Smith M, Cois A, Adjaye-Gbewonyo K*
*Estimating Trends in Cardiovascular Disease Risk for the EXPOSE (Explaining Population Trends in Cardiovascular Risk: A Comparative Analysis of Health Transitions in South Africa and England) Study: Repeated Cross-Sectional Study*
*JMIR Cardio 2025;9:e64893*
*URL: https://cardio.jmir.org/2025/1/e64893*
*doi:10.2196/64893*
*PMID:*

# Improving the Readability of Institutional Heart Failure–Related Patient Education Materials Using GPT-4: Observational Study

Ryan C King[1], MD; Jamil S Samaan[2], MD; Joseph Haquang[1], DO; Vishnu Bharani[1], MD; Samuel Margolis[3], BS; Nitin Srinivasan[4], BA; Yuxin Peng[5], BS; Yee Hui Yeo[2], MD, MSc; Roxana Ghashghaei[1], MD

[1]Department of Medicine, Division of Cardiology, University of California, Irvine Medical Center, 101 The City Dr S, Orange, CA, United States

[2]Department of Medicine, Karsh Division of Gastroenterology and Hepatology, Cedars-Sinai Medical Center, Los Angeles, CA, United States

[3]David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

[4]Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

[5]School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

**Corresponding Author:**
Ryan C King, MD
Department of Medicine, Division of Cardiology, University of California, Irvine Medical Center, 101 The City Dr S, Orange, CA, United States

## Abstract

**Background:** Heart failure management involves comprehensive lifestyle modifications such as daily weights, fluid and sodium restriction, and blood pressure monitoring, placing additional responsibility on patients and caregivers, with successful adherence often requiring extensive counseling and understandable patient education materials (PEMs). Prior research has shown PEMs related to cardiovascular disease often exceed the American Medical Association's fifth- to sixth-grade recommended reading level. The large language model (LLM) ChatGPT may be a useful tool for improving PEM readability.

**Objective:** We aim to assess the readability of heart failure–related PEMs from prominent cardiology institutions and evaluate GPT-4's ability to improve these metrics while maintaining accuracy and comprehensiveness.

**Methods:** A total of 143 heart failure–related PEMs were collected from the websites of the top 10 institutions listed on the 2022‐2023 US News & World Report for "Best Hospitals for Cardiology, Heart & Vascular Surgery." PEMs were individually entered into GPT-4 (version updated July 20, 2023), preceded by the prompt, "Please explain the following in simpler terms." Readability was assessed using the Flesch Reading Ease score, Flesch-Kincaid Grade Level (FKGL), Gunning Fog Index, Coleman-Liau Index, Simple Measure of Gobbledygook Index, and Automated Readability Index. The accuracy and comprehensiveness of revised GPT-4 PEMs were assessed by a board-certified cardiologist.

**Results:** For 143 institutional heart failure–related PEMs analyzed, the median FKGL was 10.3 (IQR 7.9-13.1; high school sophomore) compared to 7.3 (IQR 6.1-8.5; seventh grade) for GPT-4's revised PEMs ($P$<.001). Of the 143 institutional PEMs, there were 13 (9.1%) below the sixth-grade reading level, which improved to 33 (23.1%) after revision by GPT-4 ($P$<.001). No revised GPT-4 PEMs were graded as less accurate or less comprehensive compared to institutional PEMs. A total of 33 (23.1%) GPT-4 PEMs were graded as more comprehensive.

**Conclusions:** GPT-4 significantly improved the readability of institutional heart failure–related PEMs. The model may be a promising adjunct resource in addition to care provided by a licensed health care professional for patients living with heart failure. Further rigorous testing and validation is needed to investigate its safety, efficacy, and impact on patient health literacy.

***(JMIR Cardio 2025;9:e68817)*** doi:10.2196/68817

**KEYWORDS**

patient education; heart failure; artificial intelligence; large language models; ChatGPT; GPT-4; health literacy; readability

## Introduction

Heart failure affects approximately 1% ‐ 2% of adults globally, with an estimated prevalence of 64 million people [1]. Treatment involves extensive patient adherence to lifestyle modifications such as daily weights, fluid and sodium restriction, and rigorous guideline-directed medication regimens. Altogether, these interventions attempt to prevent disease progression and hospital admissions, which drive most of the financial burden ($39.2-$60 billion) related to the disease [2]. Due to the complex degree of self-management required by patients with heart failure, improving patient education and health literacy may play a crucial role in improving outcomes [3,4].

In the United States, the average adult's reading comprehension level is approximately seventh to eighth grade proficiency [5], resulting in the American Medical Association (AMA)

recommendation of written patient education materials (PEMs) being at a fifth- to sixth-grade reading level [6]. However, a 2019 readability analysis of cardiovascular disease–related PEMs reported that the mean reading level of materials was tenth grade, comparable to that of a high school sophomore [7]. Inadequate health literacy has been associated with increased relative risk of emergency department visits, hospitalizations, and mortality for patients with heart failure [4,8], highlighting the need for accessible, readable, and high-quality PEMs.

ChatGPT is a large language model (LLM) that is gaining widespread public adoption [9]. With an increasing number of patients seeking health information online [10], the model has the potential to enhance patient health education and address the complexity of heart failure–related PEMs. As ChatGPT's acceptance and usage have increased, initial research involved evaluating the model's accuracy and reliability. Several studies have shown that ChatGPT provides appropriate, accurate, and reliable knowledge across a wide range of cardiac and noncardiac medical conditions, including heart failure [11-16]. In addition to accuracy, ChatGPT has been found to deliver more empathetic responses to real-world patient questions than physicians in online forums [17]. As prior data regarding accuracy have been promising, an emerging focus has been on investigating the readability of the model's output.

Prior studies have shown ChatGPT provides accurate and comprehensive responses to questions related to heart failure, and another demonstrated its responses were at a college reading level, highlighting the need for further assessment of the readability of GPT's outputs [12,18]. Similarly, another study examining GPT-4's responses related to amyloidosis showed that while responses were often accurate and comprehensive, the average readability of responses ranged from a grade level of 10.3 (high school sophomore) to 21.7 (beyond graduate school) [16]. We aim to expand on the previous literature by assessing the readability of heart failure–related online PEMs from renowned cardiology institutions, assessing GPT-4's ability to improve the readability of these PEMs, and comparing the accuracy and comprehensiveness between institutional PEMs and GPT-4's revised PEMs.

## Methods

### Institutional Patient Education Materials

There were 143 PEMs (Multimedia Appendix 1 and Figure 1) related to heart failure collected in July 2023 from the top 10 ranked cardiology institutions (deidentified) listed on the 2022‐2023 US News & World Report website as "Best Hospitals for Cardiology, Heart & Vascular Surgery." These PEMs include frequently asked questions (FAQs) presented as text descriptions of various aspects of heart failure such as causes, symptoms, medications, and procedures. Duplicate institutional PEMs were included since education materials varied between institutions, and readability of each PEM was the primary outcome of interest.

**Figure 1.** Diagram of institutional heart failure–related PEM curation, revised GPT-4 PEM generation, and subsequent assessment of readability, accuracy, and comprehensiveness. Created in BioRender [19]. FAQ: frequently asked question; PEM: patient education material.

### GPT-4 Response Generation

Each institution's PEMs were entered into GPT-4 (version updated July 20, 2023), preceded by the prompt, "Please explain the following in simpler terms." GPT-4 was accessed using the OpenAI website interface. Default model settings were used (temperature, max tokens, etc). The "new chat" function was used for each PEM, thus creating a new conversation without a record of prior inputs. Materials containing nontext components (images or videos) were excluded.

### Readability Assessment

The readability of institutional PEMs and GPT-4's revised PEMs were then assessed using the following validated formulas: Flesch Reading Ease (FRE) score [20], Flesch-Kincaid Grade Level (FKGL) [21], Gunning Fog Index [22], Coleman-Liau Index [23], Simple Measure of Gobbledygook (SMOG) Index [24], and Automated Readability Index [25]. The FRE score, measured on a scale of 0 to 100, indicates a text with a higher score has better ease of understanding. The remaining formulas directly translate a score into its corresponding US reading grade level, such as a score of 10 translating to a tenth-grade reading level. These metrics derive their scores from the mean length of sentences and words used in a given text. In contrast to the FRE, lower scores in the other formulas correspond to an easier level of understanding. The readability formulas were assessed using the *Textstat* library in Python (Python Software Foundation) and the *Textstat readability* package in R software (R Foundation for Statistical Computing).

### Accuracy and Comprehensiveness

Accuracy and comprehensiveness of GPT-4's revised PEMs (Multimedia Appendix 1) were assessed as secondary outcomes by an actively practicing board-certified cardiologist at a tertiary academic medical center. The reviewer was not blinded during grading. The reviewer used the following grading scale in Textbox 1 when grading the original institutional PEMs and revised GPT-4 PEMs.

**Textbox 1.** Grading scale used by reviewer.

"Compared to the institutional PEM, the GPT-4 revised PEM is":

1. Less accurate
2. Equally accurate
3. More accurate

"Compared to the institutional PEM, the GPT-4 revised PEM is":

1. Less comprehensive
2. Equally comprehensiveness
3. More comprehensive

### Statistical Analysis

Descriptive statistics are presented as medians and IQRs. Readability metrics for institutional PEMs and GPT-4's revised PEMs were compared using the Mann-Whitney *U* test. Further subanalysis was performed investigating the proportion of PEMs meeting the sixth-grade reading level recommendation by the AMA among institutional PEMs and GPT-4's revised PEMs. Statistical analysis was conducted using SPSS (version 29; IBM Corporation).

### Ethical Considerations

The data collection process in this observational study did not involve patients and did not require the deidentification or protection of data. Therefore, no institutional review board approval was sought.

## Results

### Readability Assessment

Readability analysis revealed GPT-4's revised PEMs were significantly more readable compared to institutional PEMs across all 6 metrics ($P<.001$) (Figure 2). The FRE score increased from a median institutional score of 48.6 (IQR 38.0-63.3; $P<.001$; hard-to-read text, college reading level) to 72.2 (IQR 66.2-77.5; $P<.001$; fairly easy-to-read text, seventh-grade level) after GPT-4 revision [20]. The FKGL also saw improvement, decreasing from an institutional median reading level of tenth grade (IQR 7.9-13.1; $P<.001$) to seventh grade (IQR 6.1-8.5; $P<.001$) after GPT-4 revision. Furthermore, the institutional Automated Readability Index of 11.2 (IQR 7.7-14.5; $P<.001$) improved to 8.3 (IQR 6.7-9.3; $P<.001$) after GPT-4 revision. The other readability metrics (Gunning Fog Index, Coleman-Liau Index, and SMOG Index) also showed improved scores after GPT-4 revision: 9.8 (IQR 8.5-11.1; $P<.001$), 8.9 (IQR 8.1-10.0; $P<.001$), and 9.6 (IQR 8.5-10.7; $P<.001$), respectively, compared to the median institutional scores of 13.1 (IQR 10.6-16.2), 12.3 (IQR 10.1-14.5), and 12.2 (IQR 10.3-14.6). Before GPT-4 revision, 9.1% (13/143) of institutional PEMs met the AMA's recommended sixth-grade reading level (Table 1). However, after GPT-4's revision, 23.1% (33/143) of PEMs met the sixth-grade recommendation. On average, GPT-4 revision led to a 3.6 reading grade level reduction.

An example of this simplification in reading level was seen when describing different types of heart failure. The institutional PEM described right-sided heart failure as most often resulting from left-sided heart failure due to increased pressure from the left ventricle not propelling blood to the rest of the body. However, GPT-4 provided a more basic explanation using an analogy of ventricles being small rooms and gave a more

simplified explanation of right-sided heart failure as a result of left-sided heart failure. In another example, when explaining the various causes of heart failure, one institutional PEM provided a list of etiologies such as “heart valve disease” or “coronary artery disease” without a description, compared to GPT-4, which more thoroughly described the role of each cause in relation to heart failure in simple language.

**Figure 2.** Box and whiskers plot of median readability scores across 5 metrics including Automated Readability Index, Coleman-Liau Index, Flesch-Kincaid Grade Level, Gunning Fog Index, Simple Measure of Gobbledygook (SMOG) Index for institutional and GPT-4’s revised PEMs. PEMs: patient education materials. * *P*<.05.

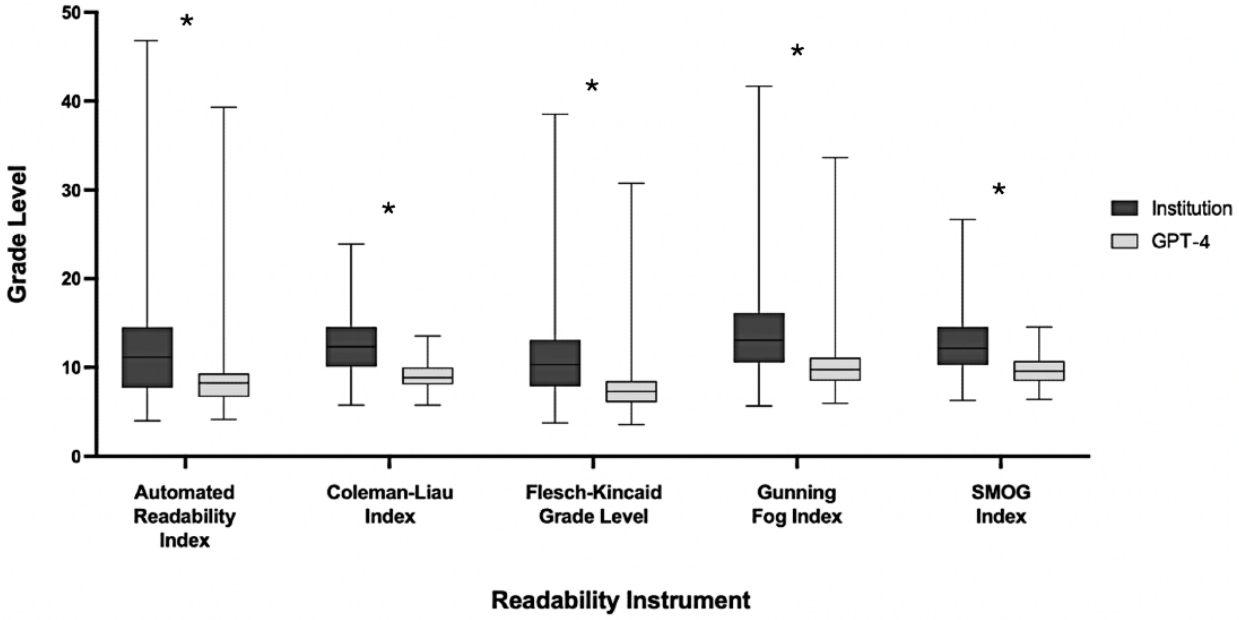



**Table .** Comparison of the proportion of patient education materials (PEMs) meeting the American Medical Association’s (AMA) recommended sixth-grade reading level between institutional and GPT-4’s revised PEMs.

| | ≤Sixth-grade reading level | ≥Sixth-grade reading level | Percent meeting AMA recommendation |
|---|---|---|---|
| Institutional Flesch-Kincaid Grade Level | 13 | 130 | 9.10 |
| GPT-4 Flesch-Kincaid Grade Level | 33 | 110 | 23.10 |

### Accuracy and Comprehensiveness

Following review by a board-certified cardiologist, 33 out of 143 (23.1%) revised GPT-4 PEMs were graded as more comprehensive than the corresponding institutional PEMs (Table 2). Additionally, all 143 (100%) revised GPT-4 PEMs were graded as equally accurate as their institutional PEM counterpart.

**Table .** Evaluation of GPT-4’s accuracy and comprehensiveness of revised patient education materials (PEMs) compared to institutional PEMs (N=143).

| Scoring | Accuracy, n (%) | Comprehensiveness, n (%) |
|---|---|---|
| Less | 0 (0) | 0 (0) |
| Equal | 143 (100) | 110 (76.9) |
| More | 0 (0) | 33 (23.1) |

## Discussion

### Principal Results

LLMs are a rapidly developing technology with the potential to enhance the delivery of PEMs to patients of all levels of health literacy. In this study, we expanded on existing research that evaluated ChatGPT’s ability to generate accurate and reliable answers to heart failure questions by examining GPT-4’s ability to improve the readability of institutional PEMs. Our analysis shows that GPT-4, when prompted, was able to significantly enhance the readability of institutional PEMs for common heart failure–related patient questions. After evaluation by a board-certified cardiologist, all of GPT-4’s revised PEMs were graded as equally accurate and many were graded as more comprehensive as institutional PEMs, with no revised PEMs graded as less accurate or less comprehensive. GPT-4’s capabilities to provide accurate, comprehensive, and readable PEMs in real-time and in a conversational manner underscores the future potential of LLMs to enhance patient education and ultimately patient health literacy.

## Comparison With Prior Work

Previous research has demonstrated that ChatGPT possesses a broad knowledge base comprising various medical conditions, including cirrhosis, hepatocellular carcinoma, and bariatric surgery [14,15,26,27]. Its knowledge base also spans cardiovascular diseases such as acute coronary syndrome [11,28], heart failure [12], atrial fibrillation [29], and even rare disorders like amyloidosis [16]—a multisystemic infiltrative disease. Specifically, regarding amyloidosis, while GPT-4 provided accurate, comprehensive, and reliable answers to gastrointestinal, neurologic, and cardiology queries, the average FKGL of responses was 15.5 (college level), significantly exceeding the recommended sixth-grade reading level set forth by the AMA [16]. Similar results were shown when examining responses to the surgical treatment of retinal diseases and hypothyroidism in pregnancy [30,31].

A previous study examined ChatGPT's ability to simplify the readability of responses to bariatric surgery–related FAQs [32]. GPT-4 reduced the average grade reading level of PEMs from eleventh (high school junior) to sixth grade, aligning with the AMA's recommendation. Another study also showed that GPT-4 improved the readability of cardiovascular magnetic resonance reports, reducing the average reading level from tenth grade to fifth grade while maintaining high factual accuracy [33]. When simplifying PEMs relating to aortic stenosis, GPT-3.5 was able to lower the mean FKGL from 9.2 to 5.9 when instructed to "translate to a 5th grade reading level" [34]. Our study further contributes to this body of work by demonstrating GPT-4's ability to improve the median readability of institutional PEMs from 10.3 (high school sophomore) to 7.3 (seventh grade) while maintaining accuracy and often enhancing comprehensiveness (Table 1). However, a unique aspect of our study was the use of a general prompt, "Please explain the following in simpler terms," compared to other studies that specifically requested simplification to a fifth- to sixth-grade reading level [34]. Our prompt simulates an organic patient encounter with the GPT-4 platform written in language meant to mirror an actual patient request for simplification. This difference in prompting but similar significant improvement in readability shows the adaptability of LLMs in this domain and may increase the likelihood of future adoption. Furthermore, the enhanced readability underscores the potential of LLMs in fostering better patient understanding of heart failure–related information.

## Limitations and Ethical Concerns

ChatGPT, while adept at generating conversational answers, has inherent limitations in accuracy and privacy. The model cannot access real-time patient records and often does not cite peer-reviewed articles or reference updated guidelines, which is crucial for accurate and evidence-based responses. Additionally, the current model may not reliably understand nuanced medical topics or accurately interpret complex medical questions [35], leading to potential patient misunderstandings. In some cases, ChatGPT may also generate answers that initially seem factual due to its confident-appearing language but disseminate inaccurate information, known as artificial hallucinations [36]. Utilizing artificial intelligence (AI) models like ChatGPT in health care settings may also not guarantee secure handling of patient information as the model may collect users' conversation data for future training. Although OpenAI does have a privacy setting allowing for disabling user data collection, prioritizing patient confidentiality will be an important aspect of development if the technology is to be used as an adjunct health care tool [37].

Furthermore, ChatGPT may also perpetuate social disparities due to implicit biases and contribute to accessibility gaps. Recent studies revealed that GPT-4 tended to promote outdated race-based medicine and overrepresent or underrepresent certain racial groups and sexes depending on the circumstance and thus potentially reinforce stereotypes [38,39]. Another concern is equitable access, as patients with lower socioeconomic status often have less access to certain technology such as the internet and may have barriers to utilizing these new AI tools [40]. Altogether, these validity and ethical considerations emphasize that clinical oversight, such as US Food and Drug Administration regulation, is warranted prior to LLM incorporation in patient care [41]. This would allow for consistent monitoring of this rapidly evolving technology, ensuring optimization of safety protocols with each new update of the model.

Our study has several limitations. Although we employed validated readability scoring systems as a surrogate for patient understanding, these formulas have their limitations, as previously reported [42,43]. These formulas often generate a reading level score that inherently grades longer words and sentences as being more complex but are unable to assess a text's content for structure and clarity. Our study also did not involve patients, which is essential for the comprehensive assessment of ChatGPT as a patient educational resource. Future studies would benefit from involving patients to ensure relevance of questions, preference in language used, and assessment of patient understanding. A baseline assessment of a patient's understanding of the given topic would also be beneficial to assess if ChatGPT can improve comprehension rather than relying on scoring tools. Additionally, we employed only one expert reviewer to assess the accuracy and comprehensiveness of ChatGPT's responses. To limit the potential for bias through subjective review and promote diverse perspectives, future research would benefit from involving multiple reviewers from different backgrounds and training institutions. Our reviewer was also not blinded to the source of each PEM, allowing for possible bias when evaluating accuracy and comprehensiveness. Our study could also not incorporate or interpret questions containing multimedia at the time of data collection, but with the release of multimodal LLMs, like GPT-4v, including visual aids would be another valuable component of PEMs to investigate. The PEMs used are not comprehensive of all questions that may be asked by patients, which limits the generalizability of our results. Future studies using real-world patients and questions would be helpful to further understand the broad spectrum of questions patients may ask.

## Future Directions

We opted for a pragmatic approach in designing the GPT-4 prompt used to revise institutional PEMs. Our focus was on

ensuring the prompt reflected a simple, intuitive command that patients would be likely to use in real-world scenarios. Although this method provided promising results, highlighting the versatility of GPT-4, exploring more intricate prompts may yield even more impressive outputs and functionality. We advocate further research into prompt engineering to better replicate natural conversations and offer specific instructions for generating higher-quality and personalized responses.

Medical institutions can utilize this technology by integrating ChatGPT directly into their online patient education platforms with customized readability based on the highest level of education completed by the patient. This type of personalization of readability assessment can be implemented in all patient-facing AI applications to ensure the appropriate reading level of text for all patients. For example, Buoy Health, a chatbot developed by Harvard Medical School in 2014, uses natural language processing to help users assess symptoms with reported accuracy rates of 90% - 98% [44,45]. Boston Children's Hospital has adopted this platform on their website to guide patients on symptoms and recommended next steps in seeking medical care [44,45]. While not solely focused on education, it demonstrates how leading institutions are successfully leveraging chatbots as interactive tools. The consideration of readability assessment and adaptability in these patient-facing applications may increase patient engagement and ensure patients of all education levels can use these tools. Greater collaboration between trusted medical institutions and LLM platforms could improve patient access to simplified, accurate medical information that aligns with the AMAs recommended fifth- to sixth-grade reading level.

## Conclusions

Our study demonstrates GPT-4's ability to improve the readability of institutional heart failure–related PEMs while also maintaining accuracy and comprehensiveness. Our results underscore the potential future utility of LLMs in improving the delivery of easy-to-understand and readable PEMs to patients of all health literacy levels. While ChatGPT may potentially be a valuable future tool in patient care, it should be used as a supplement to, rather than a replacement for, human expertise and judgment of a licensed health care professional. We recommend the development of future studies examining the optimization of readability outputs, personalization, and real-world implementation.

## Acknowledgments

ChatGPT-4 (version updated 16 May 2024), by OpenAI was used to improve readability. There was no funding obtained for this study.

## Data Availability

All data generated or analyzed during this study are included in this paper's main text and Multimedia Appendix 2.

## Conflicts of Interest

RG is a consultant for Pfizer, Alnylam, and AstraZeneca. None of the other authors have interests to disclose.

Multimedia Appendix 1
Accuracy and comprehensiveness data.
[XLSX File, 116 KB - cardio_v9i1e68817_app1.xlsx ]

Multimedia Appendix 2
Comparison of readability of institutional and GPT-4's revised patient education materials.
[PNG File, 144 KB - cardio_v9i1e68817_app2.png ]

## References

1. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. Eur J Heart Fail 2020 Aug;22(8):1342-1356. [doi: 10.1002/ejhf.1858] [Medline: 32483830]
2. Urbich M, Globe G, Pantiri K, et al. A systematic review of medical costs associated with heart failure in the USA (2014-2020). Pharmacoeconomics 2020 Nov;38(11):1219-1236. [doi: 10.1007/s40273-020-00952-0] [Medline: 32812149]
3. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011 Jul 19;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
4. Peterson PN, Shetterly SM, Clarke CL, et al. Health literacy and outcomes among patients with heart failure. JAMA 2011 Apr 27;305(16):1695-1701. [doi: 10.1001/jama.2011.512] [Medline: 21521851]
5. Fast facts: adult literacy. NCES. 2019. URL: https://nces.ed.gov/fastfacts/display.asp?id=69 [accessed 2024-10-29]
6. Weiss BD. Health Literacy: A Manual for Clinicians: American Medical Association Foundation and American Medical Association; 2003.

7. Ayyaswami V, Padmanabhan D, Patel M, et al. A readability analysis of online cardiovascular disease-related health education materials. Health Lit Res Pract 2019 Apr;3(2):e74-e80. [doi: 10.3928/24748307-20190306-03] [Medline: 31049489]
8. Fabbri M, Murad MH, Wennberg AM, et al. Health literacy and outcomes among patients with heart failure: a systematic review and meta-analysis. JACC Heart Fail 2020 Jun;8(6):451-460. [doi: 10.1016/j.jchf.2019.11.007] [Medline: 32466837]
9. Sidoti O, McClain C. 34% of U.S. adults have used ChatGPT, about double the share in 2023. Pew Research Center. 2025 Jun 25. URL: https://www.pewresearch.org/short-reads/2025/06/25/34-of-us-adults-have-used-chatgpt-about-double-the-share-in-2023/ [accessed 2025-06-26]
10. The social life of health information. Pew Research Center. 2009. URL: https://www.pewresearch.org/internet/2009/06/11/the-social-life-of-health-information [accessed 2024-10-29]
11. Sarraju A, Bruemmer D, Van Iterson E, Cho L, Rodriguez F, Laffin L. Appropriateness of cardiovascular disease prevention recommendations obtained from a popular online chat-based artificial intelligence model. JAMA 2023 Mar 14;329(10):842-844. [doi: 10.1001/jama.2023.1044] [Medline: 36735264]
12. King RC, Samaan JS, Yeo YH, Mody B, Lombardo DM, Ghashghaei R. Appropriateness of ChatGPT in answering heart failure related questions. Heart Lung Circ 2024 Sep;33(9):1314-1318. [doi: 10.1016/j.hlc.2024.03.005] [Medline: 38821760]
13. King RC, Bharani V, Shah K, Yeo YH, Samaan JS. GPT-4V passes the BLS and ACLS examinations: an analysis of GPT-4V's image recognition capabilities. Resuscitation 2024 Feb;195:110106. [doi: 10.1016/j.resuscitation.2023.110106] [Medline: 38160904]
14. Yeo YH, Samaan JS, Ng WH, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. Clin Mol Hepatol 2023 Jul;29(3):721-732. [doi: 10.3350/cmh.2023.0089] [Medline: 36946005]
15. Samaan JS, Yeo YH, Rajeev N, et al. Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery. OBES SURG 2023 Jun;33(6):1790-1796. [doi: 10.1007/s11695-023-06603-5] [Medline: 37106269]
16. King RC, Samaan JS, Yeo YH, et al. A multidisciplinary assessment of chatgpt's knowledge of amyloidosis: observational study. JMIR Cardio 2024 Apr 19;8:e53421. [doi: 10.2196/53421] [Medline: 38640472]
17. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 1;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
18. Riddell CW, Chan C, McGrinder H, Earle NJ, Poppe KK, Doughty RN. College-level reading is required to understand ChatGPT's answers to lay questions relating to heart failure. Eur J Heart Fail 2023 Dec;25(12):2336-2337. [doi: 10.1002/ejhf.3083] [Medline: 37964183]
19. King R. Figure 1. BioRender. URL: https://BioRender.com/imijjhx [accessed 2025-06-27]
20. Flesch R. Guide to academic writing. University of Canterbury School of Business and Economics. 2016. URL: https://web.archive.org/web/20160712094308/http://www.mang.canterbury.ac.nz/writing_guide/writing/flesch.shtml [accessed 2024-10-29]
21. Kincaid J, Fishburne R, Rogers R, Chissom B. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. : Institute for Simulation and Training; 1975 URL: https://stars.library.ucf.edu/cgi/viewcontent.cgi?article=1055&context=istlibrary [accessed 2025-06-25]
22. Gunning R. The Fog Index after twenty years. Journal of Business Communication 1969 Jan;6(2):3-13. [doi: 10.1177/002194366900600202]
23. Coleman M, Liau TL. A computer readability formula designed for machine scoring. Journal of Applied Psychology 1975;60(2):283-284. [doi: 10.1037/h0076540]
24. McLaughlin GH. SMOG grading: a new readability formula. J Read 1969;12(8):639-646 [FREE Full text]
25. Smith EA, Senter RJ. Automated readability index. AMRL TR 1967 May:1-14. [Medline: 5302480]
26. Samaan JS, Yeo YH, Ng WH, et al. ChatGPT's ability to comprehend and answer cirrhosis related questions in Arabic. Arab J Gastroenterol 2023 Aug;24(3):145-148. [doi: 10.1016/j.ajg.2023.08.001] [Medline: 37673708]
27. OpenAI, Achiam J, Adler S, et al. GPT-4 technical report. arXiv. Preprint posted online on Mar 15, 2023. [doi: 10.48550/arXiv.2303.08774]
28. Gurbuz DC, Varis E. Is ChatGPT knowledgeable of acute coronary syndromes and pertinent European Society of Cardiology Guidelines? Minerva Cardiol Angiol 2024 Jun;72(3):299-303. [doi: 10.23736/S2724-5683.24.06517-7] [Medline: 38391252]
29. Lee TJ, Campbell DJ, Rao AK, et al. Evaluating ChatGPT responses on atrial fibrillation for patient education. Cureus 2024 Jun;16(6):e61680. [doi: 10.7759/cureus.61680] [Medline: 38841294]
30. Onder CE, Koc G, Gokbulut P, Taskaldiran I, Kuskonmaz SM. Evaluation of the reliability and readability of ChatGPT-4 responses regarding hypothyroidism during pregnancy. Sci Rep 2024 Jan 2;14(1):243. [doi: 10.1038/s41598-023-50884-w] [Medline: 38167988]
31. Momenaei B, Wakabayashi T, Shahlaee A, et al. Appropriateness and readability of ChatGPT-4-generated responses for surgical treatment of retinal diseases. Ophthalmol Retina 2023 Oct;7(10):862-868. [doi: 10.1016/j.oret.2023.05.022] [Medline: 37277096]

32. Srinivasan N, Samaan JS, Rajeev ND, Kanu MU, Yeo YH, Samakar K. Large language models and bariatric surgery patient education: a comparative readability analysis of GPT-3.5, GPT-4, Bard, and online institutional resources. Surg Endosc 2024 May;38(5):2522-2532. [doi: 10.1007/s00464-024-10720-2] [Medline: 38472531]
33. Salam B, Kravchenko D, Nowak S, et al. Generative Pre-trained Transformer 4 makes cardiovascular magnetic resonance reports easy to understand. J Cardiovasc Magn Reson 2024;26(1):101035. [doi: 10.1016/j.jocmr.2024.101035] [Medline: 38460841]
34. Rouhi AD, Ghanem YK, Yolchieva L, et al. Can artificial intelligence improve the readability of patient education materials on aortic stenosis? A pilot study. Cardiol Ther 2024 Mar;13(1):137-147. [doi: 10.1007/s40119-023-00347-0] [Medline: 38194058]
35. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. J Med Syst 2023 Mar 4;47(1):33. [doi: 10.1007/s10916-023-01925-4] [Medline: 36869927]
36. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus 2023 Feb;15(2):e35179. [doi: 10.7759/cureus.35179] [Medline: 36811129]
37. New ways to manage your data in ChatGPT. OpenAI. 2023. URL: https://openai.com/index/new-ways-to-manage-your-data-in-chatgpt [accessed 2024-10-29]
38. Zack T, Lehman E, Suzgun M, et al. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. Lancet Digit Health 2024 Jan;6(1):e12-e22. [doi: 10.1016/S2589-7500(23)00225-X] [Medline: 38123252]
39. Omiye JA, Lester JC, Spichak S, Rotemberg V, Daneshjou R. Large language models propagate race-based medicine. NPJ Digit Med 2023 Oct 20;6(1):195. [doi: 10.1038/s41746-023-00939-z] [Medline: 37864012]
40. Wang X, Sanders HM, Liu Y, et al. ChatGPT: promise and challenges for deployment in low- and middle-income countries. Lancet Reg Health West Pac 2023 Dec;41:100905. [doi: 10.1016/j.lanwpc.2023.100905] [Medline: 37731897]
41. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based Software as a Medical Device (SaMD). : Food and Drug Administration; 2019 URL: https://www.fda.gov/media/122535/download?attachment [accessed 2025-06-26]
42. Scott B. The Gunning Fog Index (or FOG) readability formula. Readability Formula. 2025. URL: https://readabilityformulas.com/the-gunnings-fog-index-or-fog-readability-formula [accessed 2024-10-29]
43. Tip 6. use caution with readability formulas for quality reports. AHRQ. 2015. URL: https://www.ahrq.gov/talkingquality/resources/writing/tip6.html [accessed 2024-10-29]
44. Buoy Health: a chatbot that helps diagnose your symptoms. Product Hunt. 2017. URL: https://www.producthunt.com/posts/buoy-health [accessed 2025-05-01]
45. Ćirković A. Evaluation of four artificial intelligence-assisted self-diagnosis apps on three diagnoses: two-year follow-up study. J Med Internet Res 2020 Dec 4;22(12):e18097. [doi: 10.2196/18097] [Medline: 33275113]


## Abbreviations

**AI:** artificial intelligence
**AMA:** American Medical Association
**FAQ:** frequently asked question
**FKGL:** Flesch-Kincaid Grade Level
**FRE:** Flesch Reading Ease score
**LLM:** large language model
**PEM:** patient education material
**SMOG:** Simple Measure of Gobbledygook


*Edited by J Rivers; submitted 15.11.24; peer-reviewed by AD Rouhi, M Nomali; revised version received 05.06.25; accepted 08.06.25; published 08.07.25.*

*Please cite as:*
*King RC, Samaan JS, Haquang J, Bharani V, Margolis S, Srinivasan N, Peng Y, Yeo YH, Ghashghaei R*
*Improving the Readability of Institutional Heart Failure–Related Patient Education Materials Using GPT-4: Observational Study*
*JMIR Cardio 2025;9:e68817*
*URL: https://cardio.jmir.org/2025/1/e68817*
*doi:10.2196/68817*

# AI HeartBot to Increase Women's Awareness and Knowledge of Heart Attacks: Nonrandomized, Quasi-Experimental Study

Yoshimi Fukuoka[1], RN, PhD; Diane Dagyong Kim[2], MA; Jingwen Zhang[2], PhD; Thomas J Hoffmann[3], PhD; Holli A DeVon[4], RN, PhD; Kenji Sagae[5], PhD

[1]Department of Physiological Nursing, University of California, San Francisco, 521 Parnassus Ave, San Francisco, CA, United States

[2]Department of Communication, University of California, Davis, Davis, CA, United States

[3]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

[4]University of California, Los Angeles, Los Angeles, CA, United States

[5]Department of Linguistics, University of California, Davis, Davis, CA, United States

**Corresponding Author:**
Yoshimi Fukuoka, RN, PhD
Department of Physiological Nursing, University of California, San Francisco, 521 Parnassus Ave, San Francisco, CA, United States

## Abstract

**Background:** Heart disease remains a leading cause of death for women in the United States, but awareness and knowledge about it are declining. Artificial intelligence (AI) chatbots have great potential to educate women.

**Objective:** This study aimed to evaluate the potential efficacy of HeartBot to increase women's awareness and knowledge of heart attack symptoms and care-seeking behavior.

**Methods:** In this nonrandomized pilot, quasi-experimental study, 92 women aged ≥25 years without a history of heart disease completed the HeartBot interaction via SMS text messaging. The study was remotely conducted from October 2023 to January 2024. HeartBot, a fully automated AI chatbot, covered 15 topics of heart attack awareness, knowledge, symptoms, and care seeking in a single session. The mean length of the HeartBot interaction was 13.0 (SD 7.80) minutes. The primary outcomes consist of four questions: (1) recognizing signs and symptoms of a heart attack, (2) telling the difference between the signs and symptoms of a heart attack, (3) calling an ambulance or dialing 911 when experiencing heart attack symptoms, and (4) getting to an emergency room within 60 minutes after the onset of symptoms of a heart attack. Women were asked to answer the 4 questions before and after the HeartBot interaction on a scale of 1 to 4, with a higher score indicating higher levels of awareness and knowledge of heart attack risks and symptoms.

**Results:** The mean age of the sample was 45.9 (SD 11.9) years. In total, 59.8% (55/92) of the sample identified as belonging to racial or ethnic minority groups. The mean length of the HeartBot interaction was 13.0 (SD 7.80) minutes. In ordinal logistic regression models, women showed significant improvements across the 4 self-reported outcomes (ie, heart attack symptoms and calling 911) even after controlling for potential confounding factors (outcome 1: adjusted odds ratio [aOR] 7.10, 95% CI 3.52 - 13.16; outcome 2: aOR 5.47, 95% CI 2.77 - 10.78; outcome 3: aOR 5.75, 95% CI 2.86 - 11.59; and outcome 4: aOR 2.85, 95% CI 1.54 - 5.25; $P<.001$ for all 4 outcomes).

**Conclusions:** HeartBot led to a substantial increase in awareness and knowledge of heart attack risks and symptoms in women. These findings suggest that HeartBot is a promising approach to improving heart health education. A randomized controlled trial of HeartBot is warranted to establish its efficacy and safety for the clinical setting.

***(JMIR Cardio 2025;9:e80407)*** doi:10.2196/80407

**KEYWORDS**

artificial intelligence; AI; chatbot; natural language processing; heart disease; heart attack symptoms; women; conversational agent

## Introduction

### Background

Artificial intelligence (AI) chatbots using natural language processing and machine learning facilitate natural human-machine conversations. In recent years, AI chatbots have gained significant popularity in health care and research, and researchers have investigated their efficacy across health domains. AI chatbots have shown potential to improve mental health [1-3] and other chronic illnesses [4] and to promote healthy lifestyles and self-care behaviors [5-10]. However,

research on the use of chatbots in cardiovascular health is still in its infancy.

Heart disease continues to be the leading cause of mortality and morbidity for women in the United States [11]. More than 60 million women in the United States are living with heart disease [12]. Over the past 2 decades, several awareness campaigns, such as *Go Red for Women* [13] by the American Heart Association, have been conducted to educate the public and women regarding heart disease. Despite these large-scale public health campaigns, awareness of heart disease as the leading cause of death among women declined from 65% in 2009 to 44% in 2019 [14]. The greatest declines in awareness were observed among Hispanic and Black women as well as younger women [15-17].

### Objectives

We therefore need a new approach to increase knowledge and awareness of heart disease in women. We recently developed and tested an AI chatbot (hereafter referred to as "HeartBot") in a series of studies to achieve this goal. Given the rapid advancement of AI technologies and the increasing prevalence of smartphone ownership [18], HeartBot could have significant advantages over traditional public health campaigns, expanding reach and delivering personalized communication. Therefore, this study aims to evaluate the potential efficacy of a fully automated AI HeartBot in increasing women's awareness and knowledge of heart attack symptoms and care-seeking behavior.

## Methods

### Study Design and Sample

We remotely conducted a pilot, quasi-experimental study with 92 participants using the user-centered design approach [19]. Eligibility criteria were women aged ≥25 years, no self-reported cognitive impairment or history of heart disease or stroke, not a health care professional or student, not working in the health care field, living in the United States, possessed a cell phone with the ability to send and receive text messages, and had internet access. Participants were recruited via social media (ie, Facebook and Instagram [Meta]). The study was conducted from October 2023 to January 2024.

### Description of the HeartBot Platform

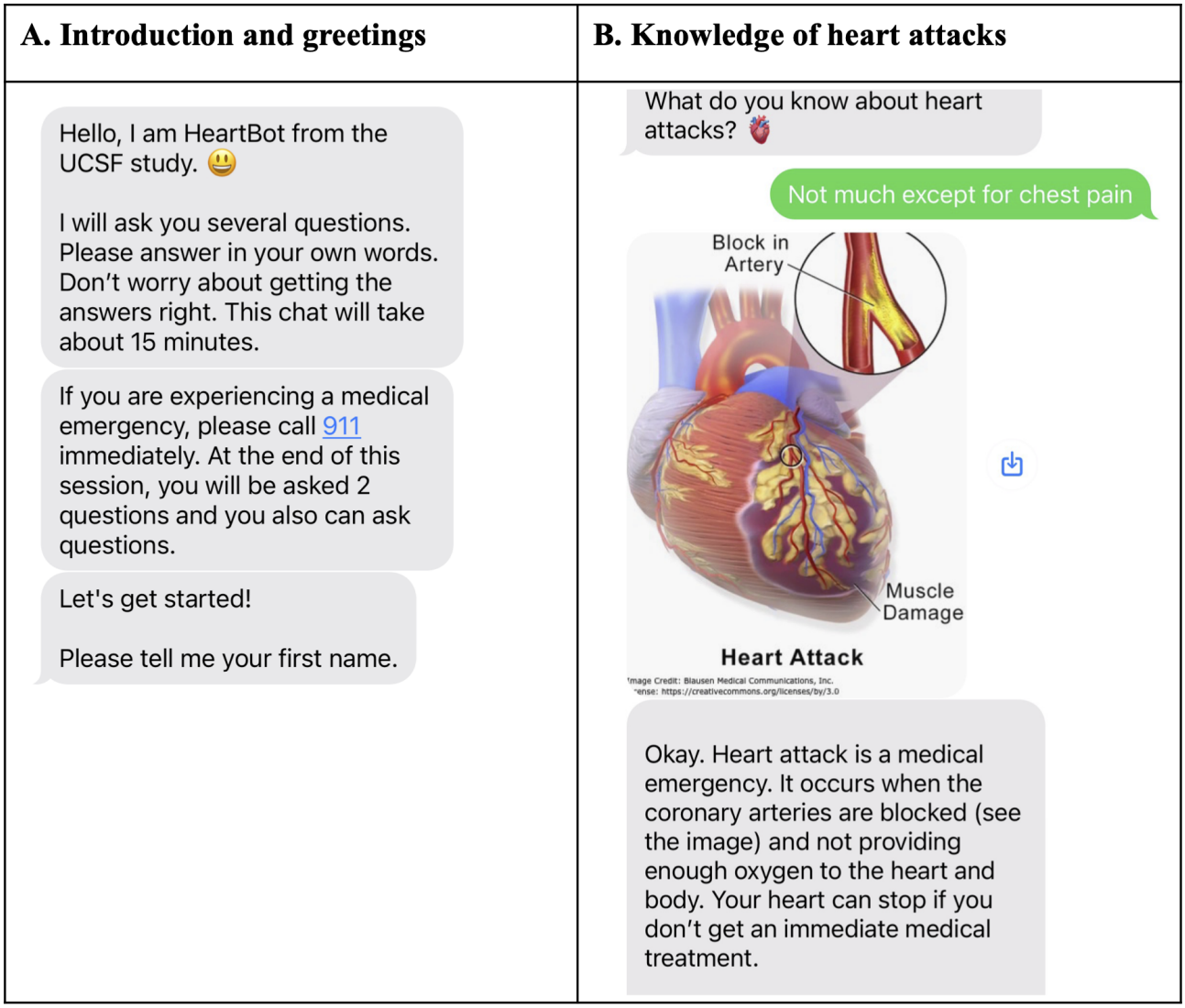
Table 1 provides an overview of the HeartBot conversation content, and Figure 1 shows the screenshots of the HeartBot conversation. HeartBot conversations occurred over SMS text messaging and started with a brief introduction, informing the user about what to expect in the conversation (Figure 1A). Within 1 conversational episode, HeartBot conversed with participants on 17 content modules, covering topics such as symptoms, risk factors, and treatment of heart attacks. To prioritize participant safety, the introduction message included the following medical emergency notice: "If you are experiencing a medical emergency, please call 911 immediately." The messages sent by HeartBot (Table 1) were developed by cardiovascular experts based on the latest guidelines and evidence to ensure full control over the content presented to participants and to minimize the risk of having the system dispense false or misleading information. In addition, we incorporated personalization and empathic responses—key communication features designed to enhance participants' experience and engagement [20]. Finally, to ensure readability, the content sent by HeartBot (Table 1) was evaluated using Flesch-Kincaid readability metrics. This analysis yielded a Flesch reading ease score of 69 and a Flesch-Kincaid grade level of 6.2, indicating that the language used was accessible and comprehensible to a broad audience.

**Table .** Overview of the message sequence and conversation content delivered by HeartBot, an artificial intelligence chatbot designed to increase women's awareness and knowledge of heart attack risks and symptoms.[a]

| Message order | Topics |
|---|---|
| 1 | Introduction and greetings |
| 2 | Participants' name retrieval |
| 3 | Knowledge of heart attacks |
| 4 | Symptoms of heart attacks |
| 5 | Leading cause of death for women in the United States |
| 6 | Gender factors for heart attacks |
| 7 | First action when experiencing symptoms of a heart attack |
| 8 | Importance of calling 911 |
| 9 | Time to seek medical help |
| 10 | Treatment of heart attacks |
| 11 | Action plans while waiting for 911 |
| 12 | Risk factors for heart disease |
| 13 | Female-specific risk factors for heart disease |
| 14 | Racial and ethnic differences in women's heart disease risk |
| 15 | Multiple-choice questions |
| 16 | Further questions to ask HeartBot |
| 17 | Acknowledgment and conclusion of the conversation |

[a]Within a single conversational episode, HeartBot delivers 17 sequential modules developed by cardiovascular experts based on current clinical guidelines. Messages were evaluated for readability using Flesch-Kincaid metrics (reading ease score=69; grade level=6.2).

**Figure 1.** Screenshots of the HeartBot conversation. The text in gray bubbles represents messages sent by HeartBot, while that in the green bubble represents the responses from the users. UCSF: University of California, San Francisco.




The current version of HeartBot, deployed to research participants, is based on an established framework that has been used in a large number of chatbots across various domains over several years. HeartBot was built using the Google Dialogflow CX platform connected to Twilio for input and output over SMS text messaging. Dialogflow CX supports the development and deployment of chatbots based on the intents and entities paradigm [21]. Every user input is categorized as one of a finite number of intents (eg, greeting, correct response, and request for help), and chatbot actions, including what the chatbot says next, are mapped directly to these intents. In HeartBot specifically, this mapping was implemented using a bidirectional encoder representations from transformers neural language model [22], which is based on the transformer architecture [23]; this approach leverages a powerful data-driven method for robustly identifying user intents expressed in free-form text messages. Within each participant's utterance, one or more entities may be mentioned. Entities are words or phrases that express specific instances of a class. For example, "high blood pressure" expresses a specific instance of the class *risk factors*. Authoring a chatbot in Dialogflow CX involves the following: (1) defining the intents that the user may express when interacting with the chatbot and which entities the chatbot should recognize within user utterances; (2) listing examples of how these intents and entities can be expressed; and (3) defining the chatbot actions that should be mapped to each intent and, optionally, to mentions of an entity. While the use of a bidirectional encoder representations from transformers language model allows the system to handle natural language input flexibly and robustly, the chatbot's behavior and responses are limited to content authored specifically for HeartBot conversations. Because each piece of the interaction is encoded directly and explicitly into the chatbot, the resulting interactions are rigid and predictable, varying in response to user utterances only in the ways intended by the system's designers. On the one hand, this ensures a great level of control over what the chatbot can do, and on the other hand, the chatbot has limited flexibility and lacks any conversational ability beyond its authored content. In addition, the design of HeartBot follows a strict system initiative philosophy: every step of the interaction between the system and the user is initiated by the system. This is in contrast to the user-initiative design embodied in current generative AI chatbots, such as ChatGPT, where the system mainly reacts to or responds to user utterances, and the user is

primarily responsible for directing the conversation. HeartBot asks specific questions, and the user responds to them. While HeartBot's utterances vary based on the user's responses, for example, by confirming correct responses or offering gentle corrections, the user has very limited control over how the interaction unfolds, and the system is in control of the conversation.

HeartBot's content library and the inventory of user intents it can handle were designed and refined based on more than 171 interactions where a research staff with domain expertise played the role of the system. These interactions took place over text messages in the same way as the HeartBot interactions, except that participants were not told they were interacting with HeartBot and had no reason to believe their conversational partner was a chatbot. The research staff conducted the conversation initially designed for HeartBot in each of these interactions. The initial version of HeartBot was based on these human-human interactions and was further refined through initial system testing, resulting in the version of HeartBot used to collect the data presented here.

## Procedures

Potential participants interested in the study were asked to complete an online screening survey on Research Electronic Data Capture (REDCap; Vanderbilt University) to determine eligibility. Research staff sent an electronic consent (e-consent) form to those meeting all eligibility criteria. Participants who signed the consent form received a baseline REDCap survey. Upon completing the baseline survey, research staff provided instructions with a specific phone number for participants to send SMS text messages to and on how to initiate a conversation with HeartBot. For participants' safety, research staff closely monitored the participants' responses during the HeartBot conversations through a back-end data monitoring platform. Research staff sent an online postsurvey link to the participants 4 to 6 weeks after the HearBot interaction. Participants who completed all study requirements received a US $20 Amazon e-gift card.

## Measures

We used a previously validated set of questions to assess the potential efficacy of HeartBot in increasing the awareness and knowledge of heart attack risks and symptoms [24,25]. These items have been applied in earlier studies with women from varied backgrounds to support their relevance across diverse populations [26-28]. Participants answered four questions before and after interacting with HeartBot, using a 4-point scale, in which 1 indicated "not sure" and 4 indicated "sure": (1) How sure are you that you could recognize the signs and symptoms of a heart attack in yourself?; (2) How sure are you that you could tell the difference between the signs or symptoms of a heart attack and other medical problems?; (3) How sure are you that you could call an ambulance or dial 911 if you thought you were having a heart attack?; and (4) How sure are you that you could get to an emergency room within 60 minutes after onset of your symptoms of a heart attack? Higher scores indicate a better awareness and knowledge of heart attack risks and symptoms.

To understand how participants engage with HeartBot and identify areas of improvement for HeartBot, we examined participants' evaluations of HeartBot using the *Effectiveness Scale*, a semantic-differential scale originally developed based on previous literature [29,30]. It consists of 5 pairs of opposite adjectives (effective vs. ineffective, helpful vs. unhelpful, beneficial vs. not beneficial, adequate vs. not adequate, supportive vs. not supportive), and each pair is scored on a 7-point Likert scale. 1 indicated the negative pole (eg, "ineffective") and 7 indicated the positive pole (eg, "effective"). Cronbach's alpha for this sample was 0.94. In addition, the impression of HeartBot was assessed by using the *Anthropomorphism Scale* [31] consisting of 5 pairs of opposite adjectives (fake vs natural, machine-like vs humanlike, unconscious vs conscious, artificial vs lifelike, rigid vs adaptive). Each pair was rated on a 7-point Likert scale, 1 being the first adjective in the pair (eg, "fake") and 7 being the second adjective (eg, "natural"). The scores for each pair in the *Effectiveness Scale* and *Anthropomorphism Scale* were summed and averaged to create mean composite scores. In this current sample, both measures had excellent reliability (Cronbach $\alpha$ 0.94 and 0.91, respectively).

## Statistical Analysis

We conducted a descriptive analysis to summarize sample characteristics, including sociodemographic characteristics, cardiovascular risks, and usability outcomes. To compare pre- and postsurvey results, we used ANOVA for continuous variables and chi-square tests for categorical variables to assess distributional differences. To determine whether participants' responses to awareness and knowledge questions significantly changed after engaging with HeartBot, we first used Wilcoxon signed-rank tests. Then, to test for differences in pre- to post-intervention awareness and knowledge of heart attack risks and symptoms as outcome responses, we fit an ordinal mixed effects logistic regression model using the R (version 4.1.0; R Foundation for Statistical Computing) [32] package *ordinal* v2022.11.16 [33], adjusting for fixed effects of White (vs non-White), age, education, income, family history, past chatbot use, mean text message effectiveness, mean HeartBot impression score, word count, conversation duration, and whether the individual thought they were texting an AI agent, as well as a random effect for individual, across the 92 participants with outcomes measured at the second time point. As a sensitivity analysis, although this model appeared to fit the data well, we also conducted a backward stepwise regression analysis ($P$<.05; resulting in 2 - 5 covariates per outcome) to ensure the model was not overparameterized. Additional sensitivity analyses included using all participants through multiple imputation for missing outcomes (instead of just imputing missing covariates), as well as a complete case analysis (2 of the 92 participants were missing income data). We also attempted a sensitivity analysis by fitting a mixed-effects multinomial logistic regression model using the generalized structural equations command in Stata (version 16.1; StataCorp LLC) [34]. However, these models did not converge, likely due to the limited sample size and increased number of parameters to estimate compared to an ordinal logistic regression model. To appropriately handle missing data, we used multiple imputation with chained

equations with 100 imputations, combined via the usual Rubin rules, using a logistic regression model for dichotomous variables and perfect mean matching for continuous variables, using the R package *mice* v3.16.0 [35]; 2 random seeds were run to ensure stability of results. All analyses used 2-tailed tests, and statistical significance was evaluated at $P<.05$. The awareness and knowledge of heart attack risks and symptoms questions were coded as time-dependent variables at baseline and post-HeartBot interaction.

### Ethical Considerations

This study was conducted in accordance with the ethical standards outlined in the Declaration of Helsinki. Institutional review board approval was obtained from the University of California, San Francisco (approval: 23‐39793). All participants provided written informed consent before study enrollment. Participation was voluntary, and participants were informed that they could withdraw at any time without penalty. All collected data were deidentified before analysis, and no personally identifiable information was retained. Data were stored on secure, password-protected servers accessible only to the research team. Participants who completed all study requirements received a US $20 Amazon e-gift card as compensation.

## Results

### Baseline Sample Characteristics

A total of 104 women completed the online baseline survey. Of these 104 women, 5 (4.8%) did not start the HeartBot interaction, and 7 (6.7%) did not complete the post-REDCap online survey. The sample characteristics of 12 women who did not complete the study did not differ from those who completed the study ($P>.05$). Table 2 presents the baseline characteristics of 92 women who completed all study requirements. The mean age was 45.9 (SD 11.9; range: 26-70) years. Out of 92 participants, 55 (59.8%) identified as belonging to racial or ethnic minority groups; 53 (57.6%) had used a chatbot such as Alexa, Google Assistant, and Siri, at least once in the past 30 days; and 13 (14.1%) reported a family history of heart attack.

**Table .** Baseline sociodemographic characteristics and cardiovascular risk factors of study participants (N=92).

| Sociodemographic characteristics | Value |
|---|---|
| Age (years), mean (SD) | 45.9 (11.9) |
| Race or ethnicity, n (%) | |
| American Indian | 1 (1.1) |
| Asian | 6 (6.5) |
| Black (non-Hispanic) | 22 (23.9) |
| Hispanic | 19 (20.7) |
| Native Hawaiian | 2 (2.2) |
| White (non-Hispanic) | 37 (40.2) |
| More than 1 race or ethnicity | 5 (5.4) |
| Education, n (%) | |
| Less than high school or did not complete college | 26 (28.3) |
| Completed college or graduate school | 66 (71.7) |
| Household income (US $), n (%) | |
| Less than 40,000 | 21 (22.9) |
| 40,001-75,000 | 30 (32.6) |
| Greater than 75,000 | 39 (42.4) |
| Do not know or decline to answer | 2 (2.2) |
| Marital status, n (%) | |
| Never married | 21 (22.8) |
| Currently married or cohabitating | 59 (64.1) |
| Divorced or widowed | 12 (13) |
| Employment status, n (%) | |
| Full time or part time | 56 (60.9) |
| Unemployed, homemaker, or student | 17 (18.5) |
| Retired, disabled, or other | 19 (20.7) |
| Chatbot use (eg, Amazon’s Alexa, Google Assistant, Siri, and Facebook Messenger bot) in the past 30 days, n (%) | |
| Yes | 53 (57.6) |
| No | 39 (42.4) |
| Self-reported cardiovascular risks, mean (SD) | |
| BMI ($kg/m^2$) | 29.5 (7.1) |
| Smoked at least 1 cigarette in the past 30 days, n (%) | |
| Yes | 14 (15.2) |
| No | 78 (84.8) |
| Blood pressure medication, n (%) | |
| Yes | 25 (27.2) |
| No or do not know | 67 (72.8) |
| Cholesterol medication, n (%) | |
| Yes | 29 (31.5) |
| No or do not know | 63 (68.5) |
| Diabetes medication, n (%) | |
| Yes | 12 (13) |
| No or do not know | 80 (85.9) |

| Sociodemographic characteristics | Value |
|---|---|
| Family history of heart disease or stroke, n (%) | |
| Yes | 13 (14.1) |
| No or do not know | 79 (85.9) |

### HeartBot Interaction

The mean and median duration of HeartBot interactions were 13.0 (SD 7.8) and 10.6 (IQR 8.5-13.9) minutes, respectively (Table 3). In addition, a post hoc analysis of HeartBot interaction logs showed that 11% (10/92) conversations had a minor technical issue in which participants were prompted to provide their name twice at the start of the session. This glitch did not interfere with delivering or completing the conversation, and all affected participants successfully completed the HeartBot session. At the end of the HeartBot conversation, 29% (27/92) of the participants submitted at least 1 question to HeartBot. The top two questions that the participants asked were regarding (1) heart disease risks and prevention risk reduction and (2) signs and symptoms of a heart attack and care-seeking behaviors during a heart attack. The participants rated HeartBot as highly effective (mean 5.7, SD 1.2) and as human-like and natural (mean 5.2, SD 1.2), with scores ranging from 1 to 7.

**Table .** HeartBot interaction (N=92).[a]

| Interaction metrics | Value |
|---|---|
| Total word count, mean (SD) | 890.1 (78.7) |
| Duration of HeartBot conversation (minutes), mean (SD) | 13.0 (7.80) |
| Number of questions participants asked, n (%) | |
| No question | 65 (70.7) |
| At least 1 question | 27 (29.3) |
| HeartBot effectiveness, mean (SD) | 5.7 (1.2) |
| HeartBot impression, mean (SD) | 5.2 (1.2) |
| “Overall, how would you rate the conversations with your texting partner?” n (%) | |
| Unnatural or very unnatural | 5 (5.4) |
| Neutral | 33 (35.9) |
| Natural or very natural | 54 (58.7) |
| “Overall, how would you rate the messages you received?” n (%) | |
| Incoherent or very incoherent | 0 (0) |
| Neutral | 23 (25.0) |
| Coherent or very coherent | 69 (75.0) |
| “Do you think you texted a human or an AI[b] chatbot during your conversation?” n (%) | |
| Human | 31 (33.7) |
| AI chatbot | 61 (66.3) |
| ”Have you ever heard or read about AI?” n (%) | |
| I would consider myself an expert in that field | 5 (5.4) |
| I could explain well and what AI is about | 27 (29.3) |
| I know somehow what AI is | 41 (44.6) |
| Yes, but I do not know exactly what AI is | 16 (17.4) |
| No | 3 (3.3) |
| “How positive or negative do you feel about the use of AI in health care?” n (%) | |
| Negative or very negative | 13 (14.1) |
| Neutral | 44 (47.8) |
| Positive or very positive | 35 (38) |
| “The use of AI will result in better health care,” n (%) | |
| Strongly disagree or disagree | 17 (18.5) |
| Neither agree nor disagree | 30 (32.6) |
| Agree or strongly agree | 45 (48.9) |
| “The use of AI will result in better health outcomes,” n (%) | |
| Disagree or strongly disagree | 16 (17.4) |
| Neither agree nor disagree | 35 (38.0) |
| Agree or strongly agree | 41 (44.6) |
| “AI may help me reduce my risk of heart disease,” n (%) | |
| Disagree or strongly disagree | 12 (13.0) |
| Neither agree nor disagree | 39 (42.4) |
| Agree or strongly agree | 41 (44.6) |

[a]Participant interaction metrics, ratings of HeartBot conversations, and attitudes toward artificial intelligence (AI): measures include conversation length, number of participant-initiated questions, perceived effectiveness, naturalness, coherence, identification of the agent as a chatbot or human, prior AI

awareness, and attitudes toward the use of AI in health care and for heart disease prevention.

[b]AI: artificial intelligence.

## Changes in Awareness and Knowledge of Heart Attack Risks and Symptoms

Table 4 shows the results of the Wilcoxon signed-rank tests for the changes in responses to awareness and knowledge of heart attack risks and symptoms questions between the pre- and post-HeartBot interaction. Participants reported significantly higher awareness and knowledge of heart attack risks and symptoms across all 4 outcome responses between the baseline and the post-HeartBot interaction ($P$<.05).

Table 5 shows the results of the ordinal mixed effects regression models for each awareness and knowledge of heart attack risks and symptoms question. Even after controlling for potential confounders (Table 5; the full model and sensitivity analysis are provided in Multimedia Appendix 1), the HeartBot interaction was significantly associated with improvements in awareness and knowledge of heart attack risks and symptoms, specifically on (1) recognizing the signs and symptoms of a heart attack response (adjusted odds ratio [aOR] 7.10, 95% CI 3.56 - 14.15; $P$<.001), (2) telling the difference between the signs or symptoms of a heart attack (aOR 5.47, 95% CI 2.77 - 10.78; $P$<.001), (3) calling an ambulance or dialing 911 during a heart attack (aOR 5.75, 95% CI 2.86 - 11.59; $P$<.001), and (4) getting to an emergency room within 60 minutes after onset of symptoms (aOR 2.85, 95% CI 1.54 - 5.25; $P$<.001). Results were very similar across all sensitivity analyses (Table S1 in Multimedia Appendix 1).

**Table .** Changes in participants' knowledge and awareness of symptoms and responses to heart attack between pre- and post-HeartBot interaction (N=92).

| | Pre-HeartBot interaction, n (%) | Post-HeartBot interaction, n (%) | $P$ value[a] |
|---|---|---|---|
| How sure are you that you could recognize the signs and symptoms of a heart attack in yourself? (please select a number from 1 - 4) | | | <.001 |
| Not sure | 24 (26.1) | 3 (3.3) | |
| Somewhat not sure | 32 (34.8) | 28 (30.4) | |
| Somewhat sure | 33 (35.9) | 40 (43.5) | |
| Sure | 3 (3.3) | 21 (22.8) | |
| How sure are you that you could tell the difference between the signs or symptoms of a heart attack and other medical problems? (please select a number from 1 - 4) | | | <.001 |
| Not sure | 28 (30.4) | 8 (8.7) | |
| Somewhat not sure | 38 (41.3) | 35 (38) | |
| Somewhat sure | 24 (26.1) | 40 (43.5) | |
| Sure | 2 (2.2) | 9 (9.8) | |
| How sure are you that you could call an ambulance or dial 911 if you thought you were having a heart attack? (please select a number from 1 - 4) | | | .02 |
| Not sure | 13 (14.1) | 3 (3.3) | |
| Somewhat not sure | 20 (21.7) | 13 (14.1) | |
| Somewhat sure | 32 (34.8) | 20 (21.7) | |
| Sure | 27 (29.3) | 56 (60.9) | |
| How sure are you that you could get to an emergency room within 60 min after the onset of your symptoms of a heart attack? (please select a number from 1 - 4) | | | <.001 |
| Not sure | 17 (18.5) | 6 (6.5) | |
| Somewhat not sure | 17 (18.5) | 12 (13) | |
| Somewhat sure | 29 (31.5) | 31 (33.7) | |
| Sure | 29 (31.5) | 43 (46.7) | |

[a]Wilcoxon signed-rank test.

**Table .** Ordinal logistic regression models for heart attack questions.

| Outcome | Unadjusted OR[a] | | Adjusted OR[b] | |
|---|---|---|---|---|
| | OR (95% CI) | *P* value | OR (95% CI) | *P* value |
| Recognize the signs and symptoms of a heart attack | 7.20 (3.63 - 14.27) | <.001 | 7.10 (3.56 - 14.15) | <.001 |
| Tell the difference between the signs or symptoms of a heart attack and other medical problems | 5.18 (2.66 - 10.09) | <.001 | 5.47 (2.77 - 10.78) | <.001 |
| Call an ambulance or dial 911 | 5.81 (2.89 - 11.68) | <.001 | 5.75 (2.86 - 11.59) | <.001 |
| Get to an emergency room within 60 minutes | 2.94 (1.61 - 5.38) | <.001 | 2.85 (1.54 - 5.25) | <.001 |

[a]OR: odds ratio.

[b]Models are adjusted for the fixed effects of race (White vs non-White), age, education, income, family history, prior chatbot use, mean text message effectiveness score, mean HeartBot impression score, word count, conversation duration, and participants' perception of texting an artificial intelligence agent, as well as a random effect for individual.

# Discussion

## Principal Findings

This study is among the first to evaluate a fully automated AI chatbot aiming to increase awareness and knowledge about heart attacks among women. All participants who started the HeartBot conversation completed it, and none left the conversation. We posit that the 3 main key factors contributed to HeartBot's success. First, HeartBot delivered a brief conversational intervention (median 10.33 min), and participants could initiate the intervention anywhere and at any time. This design dramatically reduces participant burden and ensures a potential for large-scale dissemination and effectiveness. In addition, the HeartBot conversation was delivered via SMS text messaging and did not require Wi-Fi access or a smartphone. Nearly all Americans (98%) own a mobile phone [36], and using SMS text messaging is an effective way to reach minority populations and ensure intervention equity. Second, HeartBot incorporated personalization and empathetic responses that are known to increase participants' engagement [11-13] during conversations. Addressing users by name within conversations could have fostered a sense of individual relevance and strengthened a positive evaluation of HeartBot. Third, the content of HeartBot was developed from scientific evidence and guidelines and used simple nontechnical language with images, ensuring that the educational content was easily understandable and accessible to individuals with low health literacy levels.

## Comparison With Prior Work

Previous Hispanic and Black women, and younger age groups, have shown a significant decline in awareness of heart disease as the leading cause of death among women [14-17]. The latest study reported that the overall awareness dropped from 65% in 2009 to 44% in 2019, with the steepest declines among Hispanic women (28.9%), non-Hispanic Black women (28.1%), and women aged 25 to 34 years (41.3%) [14]. To address this concern, we ensured that women with racial or ethnic minority backgrounds were well represented in this study sample, and the HeartBot included a content module specifically discussing the increased risk of heart attack in Hispanic and Black women [37,38]. HeartBot provided unbiased interactions with all participants, reducing potential biases that might arise from human facilitators in traditional in-person interventions, such as implicit judgments [39,40]. The HeartBot intervention appears to work regardless of age and race or ethnicity. To our knowledge, this is the first AI-driven chatbot intervention specifically designed to increase women's awareness and knowledge of heart disease. The greatest advantage of the HeartBot intervention over conventional public heart attack awareness campaigns is its ability to promote active learning through personalized and SMS-based dialogue. Traditional heart attack campaigns face several challenges, such as passive information delivery, limited personalization and interactivity, and dissemination gaps. In contrast, personalized chatbots can foster active learning by prompting users to engage in feedback retrieval, self-reflection, and goal setting during interactions [41,42]. We believed that the inclusion of a follow-up quiz reviewing key content at the end of the session further enhanced active learning throughout the session.

Overall, the findings of this HeartBot trial hold great promise. Although we acknowledge that generative AI models are improving at a rapid pace, we think deploying a hybrid chatbot, which combines data-driven identification of user intents with scientifically vetted conversational content, is essential to prioritize participants' safety and accuracy of information delivery. Given the purpose of the intervention to increase awareness and knowledge of heart attacks, the current design, which combines a structured content flow with personalized conversational strategies, may be optimal.

## Limitations

Several limitations of this study must be acknowledged. First, because this study was not a randomized controlled trial (RCT), causal relationships and definitive efficacy cannot be established. However, we adjusted potential confounding factors in our analysis. Second, relying on a convenience sample of women may introduce selection bias. In addition, only

participants who had at least moderate digital literacy, were comfortable with mobile technology, and were interested in their own health might be enrolled in this study, which could have influenced both engagement with the chatbot and the outcomes. To mitigate these section biases, we enrolled a wide age range (ie, 26-70 y) and a diverse sample of women (ie, 55/92, 60% belonging to racial or ethnic minority groups), as well as designed the study to not require Wi-Fi access (SMS text messaging only). Finally, the primary outcomes were assessed using self-reported measures of awareness and knowledge of heart attack risks and symptoms over a short period; behavioral changes were not assessed. Social desirability bias may have led some participants to overreport positive impressions or improvements in knowledge following the HeartBot interaction.

### Future Direction

A full-scale RCT is warranted to examine the effectiveness of the HeartBot intervention in a large, racially and ethnically diverse sample. Future research also needs to aim to assess whether increased awareness and knowledge will lead to early access to care when women are experiencing heart attack symptoms. The framework of participants' engagement and safety metrics must be developed to demonstrate use for clinical practice. Finally, future research should explore the duration of the effects of new awareness and knowledge and whether they lead to desired behaviors in an RCT. To sustain the awareness and knowledge of heart attacks, the HeartBot intervention may require multiple sessions, and the dosage and content of the intervention may need further adjustment.

### Conclusions

Interaction with HeartBot was associated with increased awareness and knowledge of heart attack risks and symptoms in a national sample of US women. These findings suggest that AI chatbot–based interventions may be a promising approach to improve women's knowledge and awareness of heart attack in the United States. An RCT of HeartBot is warranted to establish its efficacy and safety before implementation in clinical settings.

## Acknowledgments

This project was supported by the Noyce Foundation and the University of California, San Francisco School of Nursing Emile Hansen Gaine Fund. The project sponsors had no role in the study design; collection, analysis, or interpretation of the data; writing the manuscript; or the decision to submit the manuscript for publication.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

YF had full access to all study data and is responsible for the integrity and accuracy of the data analysis. YF, JZ, HAD, and KS conceptualized and designed the study. YF, DDK, JZ, HAD, and KS developed the chatbot. YF and DDK handled data acquisition, and YF, DDK, and TJH conducted the statistical analysis. YF, DDK, TJH, and KS drafted the manuscript. YF oversaw supervision and funding, while DDK provided administrative, technical, or material support. All authors critically reviewed the manuscript for important intellectual content.

## Conflicts of Interest

None declared.

Multimedia Appendix 1

Full regression results for outcome questions. The "Adjusted OR" reflects results from the subset of participants with measured outcomes (n=92), as reported in the paper. The first sensitivity analysis includes the full dataset (n=104), and the second is a complete case analysis (n=90; 2 participants were missing income data). The remaining sensitivity analyses correspond to these 3 models but use backward stepwise regression. Cells show odds ratio (95% CI). Outcomes were as follows: (1) recognizing the signs and symptoms of a heart attack, (2) distinguishing the signs or symptoms of a heart attack from other medical problems, (3) calling an ambulance or dialing 911, and (4) reaching an emergency room within 60 minutes.
[XLSX File, 14 KB - cardio_v9i1e80407_app1.xlsx ]

## References

1. Zhong W, Luo J, Zhang H. The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: a systematic review and meta-analysis. J Affect Disord 2024 Jul 1;356:459-469. [doi: 10.1016/j.jad.2024.04.057] [Medline: 38631422]
2. He Y, Yang L, Qian C, et al. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2023 Apr 28;25:e43862. [doi: 10.2196/43862] [Medline: 37115595]

3. Lim SM, Shiau CW, Cheng LJ, Lau Y. Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: a systematic review and meta-regression. Behav Ther 2022 Mar;53(2):334-347. [doi: 10.1016/j.beth.2021.09.007] [Medline: 35227408]
4. Kurniawan MH, Handiyani H, Nuraini T, Hariyati RT, Sutrisno S. A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness. Ann Med 2024 Dec;56(1):2302980. [doi: 10.1080/07853890.2024.2302980] [Medline: 38466897]
5. Oh YJ, Zhang J, Fang ML, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. Int J Behav Nutr Phys Act 2021 Dec 11;18(1):160. [doi: 10.1186/s12966-021-01224-6] [Medline: 34895247]
6. Noh E, Won J, Jo S, Hahm DH, Lee H. Conversational agents for body weight management: systematic review. J Med Internet Res 2023 May 26;25:e42238. [doi: 10.2196/42238] [Medline: 37234029]
7. Bendotti H, Lawler S, Chan GC, Gartner C, Ireland D, Marshall HM. Conversational artificial intelligence interventions to support smoking cessation: a systematic review and meta-analysis. Digit Health 2023 Nov 3;9:20552076231211634. [doi: 10.1177/20552076231211634] [Medline: 37928336]
8. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. J Med Internet Res 2023 Feb 24;25:e40789. [doi: 10.2196/40789] [Medline: 36826990]
9. Kim HK. The effects of artificial intelligence chatbots on women's health: a systematic review and meta-analysis. Healthcare (Basel) 2024 Feb 23;12(5):534. [doi: 10.3390/healthcare12050534] [Medline: 38470645]
10. Lyzwinski LN, Elgendi M, Menon C. Conversational agents and avatars for cardiometabolic risk factors and lifestyle-related behaviors: scoping review. JMIR Mhealth Uhealth 2023 May 25;11:e39649. [doi: 10.2196/39649] [Medline: 37227765]
11. Benjamin EJ, Muntner P, Alonso A, et al. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. Circulation 2019 Mar 5;139(10):e56-e528. [doi: 10.1161/CIR.0000000000000659] [Medline: 30700139]
12. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation 2022 Feb 22;145(8):e153-e639. [doi: 10.1161/CIR.0000000000001052] [Medline: 35078371]
13. Go Red for Women. American Heart Association. URL: https://www.goredforwomen.org/en/ [accessed 2025-09-23]
14. Cushman M, Shay CM, Howard VJ, et al. Ten-year differences in women's awareness related to coronary heart disease: results of the 2019 American Heart Association national survey: a special report from the American Heart Association. Circulation 2021 Feb 16;143(7):e239-e248. [doi: 10.1161/CIR.0000000000000907] [Medline: 32954796]
15. Mosca L, Hammond G, Mochari-Greenberger H, Towfighi A, Albert MA. Fifteen-year trends in awareness of heart disease in women: results of a 2012 American Heart Association national survey. Circulation 2013 Mar 19;127(11):1254-1263. [doi: 10.1161/CIR.0b013e318287cf2f] [Medline: 23429926]
16. Mosca L, Mochari-Greenberger H, Dolor RJ, Newby LK, Robb KJ. Twelve-year follow-up of American women's awareness of cardiovascular disease risk and barriers to heart health. Circ Cardiovasc Qual Outcomes 2010 Mar;3(2):120-127. [doi: 10.1161/CIRCOUTCOMES.109.915538] [Medline: 20147489]
17. Mosca L, Ferris A, Fabunmi R, Robertson RM, American Heart Association. Tracking women's awareness of heart disease: an American Heart Association national study. Circulation 2004 Feb 10;109(5):573-579. [doi: 10.1161/01.CIR.0000115222.69428.C9] [Medline: 14761901]
18. Gelles-Watnick R. Americans' use of mobile technology and home broadband. Pew Research Center. 2024. URL: http://www.jstor.org/stable/resrep63511 [accessed 2025-09-23]
19. Norman DA, Draper SW. User-Centered System Design: New Perspectives on Human-Computer Interaction: Lawrence Erlbaum Associates; 1986. [doi: 10.1201/9780367807320]
20. Adikari A, de Silva D, Moraliyage H, et al. Empathic conversational agents for real-time monitoring and co-facilitation of patient-centered healthcare. Fut Gen Comput Syst 2022 Jan 1;126(C):318-329. [doi: 10.1016/j.future.2021.08.015]
21. Williams JD, Kamal E, Ashour M, Amr H, Miller J, Zweig G. Fast and easy language understanding for dialog systems with microsoft language understanding intelligent service (LUIS). Presented at: Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue; Sep 2-4, 2015; Prague, Czech Republic URL: http://aclweb.org/anthology/W15-46 [accessed 2025-09-23] [doi: 10.18653/v1/W15-4622]
22. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. arXiv. Preprint posted online on Jun 12, 2017. [doi: 10.48550/arXiv.1706.03762]
23. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. Presented at: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; Jun 2-7, 2019; Minneapolis, MN. [doi: 10.18653/v1/N19-1423]
24. Luepker RV, Raczynski JM, Osganian S, et al. Effect of a community intervention on patient delay and emergency medical service use in acute coronary heart disease: the Rapid Early Action for Coronary Treatment (REACT) trial. JAMA 2000 Jul 5;284(1):60-67. [doi: 10.1001/jama.284.1.60] [Medline: 10872014]
25. Riegel B, McKinley S, Moser DK, Meischke H, Doering L, Dracup K. Psychometric evaluation of the acute coronary syndrome (ACS) response index. Res Nurs Health 2007 Dec;30(6):584-594. [doi: 10.1002/nur.20213] [Medline: 18022812]

26. Fukuoka Y, Lisha NE, Vittinghoff E. Comparing Asian American women's knowledge, self-efficacy, and perceived risk of heart attack to other racial and ethnic groups: the mPED trial. J Womens Health (Larchmt) 2017 Sep;26(9):1012-1019. [doi: 10.1089/jwh.2016.6156] [Medline: 28418750]
27. Fukuoka Y, Oh YJ. Perceived heart attack likelihood in adults with a high diabetes risk. Heart Lung 2022;52:42-47. [doi: 10.1016/j.hrtlng.2021.11.007] [Medline: 34856501]
28. Fukuoka Y, Oh YJ. Perceived risk of heart attack and type 2 diabetes in Hispanic adults with overweight and obesity. J Cardiovasc Nurs 2022;37(6):E197-E205. [doi: 10.1097/JCN.0000000000000866] [Medline: 34740229]
29. Liao W, Oh YJ, Feng B, Zhang J. Understanding the influence discrepancy between human and artificial agent in advice interactions: the role of stereotypical perception of agency. Commun Res 2023;50(5):633-664. [doi: 10.1177/00936502221138427]
30. Feng B. Testing an integrated model of advice giving in supportive interactions. Hum Commun Res 2009 Jan;35(1):115-129. [doi: 10.1111/j.1468-2958.2008.01340.x]
31. Bartneck C, Kulić D, Croft E, Zoghbi S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J Soc Robot 2008 Nov;1:71-81. [doi: 10.1007/s12369-008-0001-3]
32. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2010. URL: https://www.R-project.org [accessed 2025-09-23]
33. Christensen RH. Ordinal - regression models for ordinal data. The Comprehensive R Archive Network. 2024. URL: https://cran.r-project.org/web/packages/ordinal/index.html [accessed 2025-09-23]
34. Stata statistical software: release 16. StataCorp LLC. 2019. URL: https://www.stata.com/ [accessed 2025-09-23]
35. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
36. Mobile fact sheet. Pew Research Center. 2024 Nov 13. URL: https://www.pewresearch.org/internet/fact-sheet/mobile [accessed 2025-09-23]
37. Blackston JW, Safford MM, Mefford MT, et al. Cardiovascular disease events and mortality after myocardial infarction among Black and White adults: REGARDS study. Circ Cardiovasc Qual Outcomes 2020 Dec;13(12):e006683. [doi: 10.1161/CIRCOUTCOMES.120.006683] [Medline: 33302710]
38. Ya'qoub L, Lemor A, Dabbagh M, et al. Racial, ethnic, and sex disparities in patients with STEMI and cardiogenic shock. JACC Cardiovasc Interv 2021 Mar 22;14(6):653-660. [doi: 10.1016/j.jcin.2021.01.003] [Medline: 33736772]
39. Chapman EN, Kaatz A, Carnes M. Physicians and implicit bias: how doctors may unwittingly perpetuate health care disparities. J Gen Intern Med 2013 Nov;28(11):1504-1510. [doi: 10.1007/s11606-013-2441-1] [Medline: 23576243]
40. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. BMC Med Ethics 2017 Mar 1;18(1):19. [doi: 10.1186/s12910-017-0179-8] [Medline: 28249596]
41. Lin MP, Chang D. Enhancing post-secondary writers' writing skills with a chatbot. Educ Technol Soc 2020;23(1):78-92 [FREE Full text]
42. Vanichvasin P. Chatbot development as a digital learning tool to increase students' research knowledge. Int Educ Stud 2021;14(2):44. [doi: 10.5539/ies.v14n2p44]


**Abbreviations:**

**AI:** artificial intelligence
**aOR:** adjusted odds ratio
**RCT:** randomized controlled trial
**REDCap:** Research Electronic Data Capture


*Edited by A Coristine; submitted 09.07.25; peer-reviewed by R Singh, VV Sangaraju; revised version received 19.08.25; accepted 31.08.25; published 15.10.25.*

*Please cite as:*
*Fukuoka Y, Kim DD, Zhang J, Hoffmann TJ, DeVon HA, Sagae K*
*AI HeartBot to Increase Women's Awareness and Knowledge of Heart Attacks: Nonrandomized, Quasi-Experimental Study*
*JMIR Cardio 2025;9:e80407*
*URL: https://cardio.jmir.org/2025/1/e80407*
*doi:10.2196/80407*

Original Paper

# Optimization of the Care4Today Digital Health Platform to Enhance Self-Reporting of Medication Adherence and Health Experiences in Patients With Coronary or Peripheral Artery Disease: Mixed Methods Study

Stephanie Juan[1], MPH; Ante Harxhi[1], MD; Simrati Kaul[1], PhD; Breeana Woods[2], MPH; Monica Tran[1], PharmD; Gabrielle Geonnotti[1], MPH; Archit Gupta[3], MBA; Emily Dean[4], MA; Cassandra E Saunders[5], BA; Gloria Payne[6], MEd

[1]Janssen Scientific Affairs, LLC, Titusville, NJ, United States

[2]Johnson & Johnson Technology Services, Titusville, NJ, United States

[3]ZS Associates, San Francisco, CA, United States

[4]CorEvitas, part of ThermoFisher Scientific, Waltham, MA, United States

[5]Participant in Janssen Patient Engagement Research Council, Philadelphia, PA, United States

[6]Participant in Janssen Patient Engagement Research Council, Maynard, MA, United States

**Corresponding Author:**
Stephanie Juan, MPH
Janssen Scientific Affairs, LLC
1125 Trenton-Harbourton Road
Titusville, NJ, 08560
United States
Phone: 1 609 732 5221
Email: stphjuan@gmail.com

## Abstract

**Background:** Care4Today is a digital health platform developed by Johnson & Johnson comprising a patient mobile app (Care4Today Connect), a health care provider (HCP) portal, and an educational website. It aims to improve medication adherence; enable self-reporting of health experiences; provide patient education; enhance connection with HCPs; and facilitate data and analytics learning across disease areas, including cardiovascular disease.

**Objective:** This study aimed to gather patient feedback on Care4Today Connect, specifically the coronary artery disease (CAD) and peripheral artery disease (PAD) module, and to cocreate and validate features with patients to optimize the app experience for those with CAD, PAD, or both.

**Methods:** We conducted 3 research engagements between November 2022 and May 2023. Participants were US-based adults recruited and consented through the sponsor's Patient Engagement Research Council program. Participants self-reported a diagnosis of cardiovascular disease, and in some cases, specifically, CAD, PAD, or both. Part 1, internet survey, posed quantitative questions with Likert-scale answer options about existing app features. Part 2, virtual focus group, and part 3, virtual individual interviews, both used semistructured qualitative discussion to cocreate and validate new app enhancements. The quantitative data from part 1 was evaluated descriptively to categorize mobile health app use, confidence in the ability to use the app, and motivations for app use. The qualitative discussions from parts 2 and 3 were synthesized to understand participants' app needs and preferences to inform an optimal app experience.

**Results:** The response rate for part 1, internet survey, was 67% (37/55). Most participants felt at least somewhat confident using the app after seeing the newly added app tutorial (33/37, 89%), and at least somewhat confident in their ability to earn points for completing activities using app instructions (33/37, 89%). In part 2, virtual focus group (n=3), and part 3, virtual individual interviews (n=8), participants collectively preferred to enhance the app with (1) the ability to automatically add medication data for tracking and (2) the ability to receive relevant care team feedback on their self-reported health experiences. Participants would be willing to spend 10-15 minutes a day tracking 4-5 health experiences, especially those requested by their HCP.

**Conclusions:** Participants prefer apps that can reduce user burden and provide information relevant to them. Care4Today Connect can optimize the user experience for patients with CAD, PAD, or both with the automatic addition of medication data for tracking and in-app care team feedback on patient self-reported health experiences.

***(JMIR Cardio 2025;9:e56053)*** doi:10.2196/56053

**KEYWORDS**

app; cardiovascular disease; Care4Today; coronary artery disease; digital health; health tracker; medication reminder; mobile health; mHealth; qualitative; peripheral artery disease

## Introduction

### Overview

With the widespread use of mobile health (mHealth) apps and wearable fitness trackers, many people routinely self-report personal health experiences (eg, physical activity, sleep, and mood). In the health care context, self-reported data are useful for shared decision-making, providing clinicians with a more holistic perspective of patient health beyond office visits and hospitalizations, improving communication, enhancing coordination of care, and increasing patient engagement [1]. Digital health technology has the potential to become an important part of health care systems, promoting behavior change, enhancing medication adherence, and improving health outcomes in chronic conditions such as cardiovascular disease [2].

More than 18 million adults in the United States have coronary artery disease (CAD) [3], and up to 42% of these people also have peripheral artery disease (PAD) [4,5]. CAD remains the leading cause of death in the United States, accounting for 1 in every 4 deaths [6], and medication nonadherence is linked to poor outcomes [7]. Patients with CAD, PAD, or both often take multiple medications to control their disease and other comorbid conditions. The prevalence of polypharmacy (typically defined as simultaneous use of ≥5 medications [8]) is estimated to be 17% among US adults, 40% to 62% in those with heart disease [9], and 91% in patients with CAD [10]. Polypharmacy has been linked to both medication errors [11] and nonadherence [12].

In CAD, mHealth apps have been shown to support secondary prevention lifestyle changes [13], with positive effects on medication adherence, exercise and physical activity, quality of life, major adverse cardiovascular outcomes, and hospital readmissions [14-16]. In PAD, mHealth technologies have been used successfully to improve health behavior, providing motivation to exercise through activity monitoring and coaching, and have been linked to changes in both health outcomes and disease coping [17].

### Care4Today

Care4Today is a digital health platform initially launched by Johnson & Johnson as a medication reminder app in 2013. Today, the platform has expanded to 3 components: a patient mobile app, a health care provider (HCP) portal, and an educational website. The app (Care4Today Connect [18]) has been designed to encourage patients to take an active role in managing their overall health. According to the sponsor's internal health store database and Google Analytics, from mid-2020 until mid-2024, the app has supported an estimated 2000 users across company-sponsored initiatives. Features include medication and appointment reminders; various self-reported health experience trackers, including elective biometrics, health, and lifestyle activity with visual trends over time; and educational resources tailored toward specific disease management. Users can access scheduled health activities and resources related to their disease and can share data on their progress with their care team. Access to the app is granted to users in the United States with a code provided by their HCP across multiple disease areas, including cardiovascular disease [18]. It is available for both iOS and Android users; is available in English and Spanish; and can connect to fitness apps like HealthKit, Google Fit, and Fitbit but does not require a wearable device.

The Care4Today HCP portal allows the care team to view patient self-reported health experiences shared through the mobile app. The portal enables the care team to assign, monitor, and adjust patient care (eg, medications, appointments, education, and trackers) in real time, as well as to send in-app reminders and encouragement to their patients. The Care4Today website [18] provides additional educational resources accessible to both patients and HCPs. A cardiovascular health-specific webpage was created to complement the CAD- and PAD-specific care modules for the app.

Patient cocreation and validation are essential for optimizing the mobile app experience and app usefulness for managing disease. Quantitative surveys are a valuable means of capturing patient feedback, while qualitative studies can provide rich context about patient perspectives, the user experience, and barriers to using apps for health management. We conducted a 3-part, exploratory study to optimize the Care4Today Connect app and digital health platform for people living with CAD, PAD, or both, via a mixed methods approach involving both quantitative and qualitative components.

## Methods

### Ethical Considerations

A consent and release form was signed by the participants that communicated confidentiality and Health Insurance Portability and Accountability Act (HIPAA)–compliant practices. This study (institutional review board [IRB] ID 12459-EDean) was assessed by Sterling IRB (Atlanta, GA) and determined to be exempt from IRB review (45 C.F.R. §46.104(d)) under the Department of Health and Human Services category 2 exemption. The purpose of this study was to collect personal perspectives and qualitative insights from the participants. The

study was also conducted in accordance with the Helsinki Declaration of 1964 and its later amendments. The study was voluntary, and all participants were compensated for their time.

## Study Design

This exploratory sequential research was conducted in three parts: part 1, internet survey, to gain patient feedback on existing features of the Care4Today Connect app; part 2, virtual focus group, in which participants collectively helped to cocreate and envision app enhancements; and part 3, virtual individual interviews, to validate prototype app enhancements discussed in part 2.

## Participant Recruitment

Adults with cardiovascular disease residing in the United States were recruited and consented through the sponsor's Patient Engagement Research Council (PERC) program. PERCs constitute groups of disease-aware individuals living with chronic health conditions in the United States [19,20]. People with a range of health care experiences are recruited based on clinical, demographic, and epidemiologic criteria through various channels, including outreach to patient advocacy organizations, digital advertisements, social media, and physician referrals. PERC members come together to share their experiences and insights of a common diagnosis through a structured series of specific engagement activities.

Eligible participants for all 3 parts of this study were members of the sponsor's PERC who self-reported having a diagnosed cardiovascular condition. Purposeful sampling was used to ensure racial and ethnic diversity across all parts of the study. Full eligibility criteria for PERC members are described in Multimedia Appendix 1. In part 2, purposeful sampling was used to ensure that all participants were taking >1 medication (self-reported) and that a variety of experience levels with mHealth apps was represented.

## Procedures

### *Part 1: Internet Survey*

Part 1, internet survey, was conducted with participants with cardiovascular disease, including those with CAD, PAD, or both, between November 28 and December 2, 2022. Eligible participants were invited to participate via email and received a survey link programmed using Alchemer software. CorEvitas designed the survey to be completed within 25 minutes. The aim was to assess respondent's understanding of how to use existing app features. It consisted of 33 questions across 5 categories, including Upfront, Tutorial for New Users, Earned Points, App in Clinical Study, and Overall. Three "Upfront" questions focused on the demographics and clinical characteristics of respondents, and their experience with mHealth apps. The "Tutorial for New Users" category included 18 questions asking the respondent to review tutorial screenshots of how to navigate the app as well as indicate their understanding of each component. The "Earned Points" category included 4 questions asking the respondent to review app screenshots on how to earn points for completing app activities and indicate their understanding of each component. They were also asked to share their opinions on the concept of earning points for completing activities in the app. The "App in Clinical Study" included 6 questions about motivations for taking part in a clinical study using mHealth apps (Multimedia Appendix 2). The "Overall" category included 2 questions asking the respondent to indicate how likely they would be to recommend the app to a friend or coworker. The rating scale was 1 to 10, where 1 was unlikely and 10 was very likely. For most questions, multiselect or 5-point, Likert-scale response options of agree to disagree, or not at all confident, to very confident were provided, including an option to choose "Other" and elaborate in a free-text response.

### *Part 2: Cocreation*

Part 2, virtual focus group, was held on April 13, 2023, with participants with CAD, PAD, or both. The aim was to cocreate concepts with a small group of participants. Design and facilitation were led jointly by researchers from CorEvitas and ZS Associates. During the 2-hour session, participants were given an overview of the Care4Today Connect app and were asked to discuss features that may enhance the user experience. A semistructured discussion guide focused the session on two initiatives: (1) features that could improve how medication data are added to the app to ensure correct prescribed medications are tracked, alleviate user burden of manual input, and reduce input error; and (2) features for improved sharing of self-reported health experiences that could be used to facilitate feedback from care teams. To aid discussions, additional information was shared with the group, including screenshots of the existing feature for adding medication data (Multimedia Appendix 3) and illustrative mock-ups of how new medication, as well as health experience tracking features that might be incorporated into the app (Multimedia Appendix 4). For adding medication data, 2 options were presented; option 1 leveraged third-party insurance portal while option 2 used optical character recognition (OCR) technology, which involves the user taking an image of a medication bottle and then converting that image to readable text [21]. For self-reporting of health experiences, the existing method for tracking this data was presented.

### *Part 3: Validation*

Part 3 of the research aimed to validate the enhancements cocreated with patients during the virtual focus group in part 2. One-hour virtual interviews were conducted between May 2 and 4, 2023, with participants with CAD, PAD, or both. Design and facilitation were led jointly by researchers from CorEvitas and ZS Associates. Discussions were structured around two enhancements identified in part 2: (1) auto-add medication data via the insurance portal and OCR; and (2) a "For You" tab with notifications, and personalized feedback about trends in their self-reported medication or health experiences tracking. To help with this, visuals were provided of Care4Today Connect app prototypes (Multimedia Appendix 5), and a semistructured discussion guide (Multimedia Appendix 6) was used to focus the agenda. Participants were asked to rate the perceived value of, and their willingness to use, the proposed features on a 7-point Likert scale (1=not at all likely; 7=highly likely).

### Data Collection

All participants provided insight into their current experience with medication and health experience tracking and their prior use of mHealth apps. Demographic information was also collected in the part 1 web-based survey. All sessions were audio recorded and transcribed.

### Analysis

#### *Part 1: Internet Survey*

Quantitative analysis was applied to summarize collective responses in Microsoft Excel. The goal of the analysis was to assess the user's understanding of how to use existing app features. A senior patient experience research specialist from CorEvitas reviewed and presented the data descriptively as frequency and percentage.

#### *Part 2: Focus Group and Part 3: Individual Interviews*

Qualitative analysis identified patient insights and preferences directly applicable to the Care4Today app. The goal of the analysis was to detail the recommended features to be incorporated into a future version of the app. The team of senior research specialists and product designers from ZS Associates directly observed and analyzed the data. Patient insights were synthesized by using a directed content approach where inputs were systematically mapped to potential app functionalities presented during each session. The data were then further categorized by user appeal, task ease, and privacy concerns, and then finally synthesized to inform whether to enhance, modify, or deprioritize discussed C4T enhancements. No formal coding was used.

## Results

### Overview

Figure 1 provides a visual diagram of the overall mixed methods design and participant disposition. Participant demographics for each of the 3 parts of the study are described in Table 1. Table 2 describes medication tracking and health experience reporting behavior for participants in parts 2 and 3.

**Figure 1.** Study design and participant disposition. CAD: coronary artery disease; CVD: cardiovascular disease; CVM: Cardiovascular and Metabolic; mHealth: mobile health; PAD: peripheral artery disease; PERC: Patient Engagement Research Council.

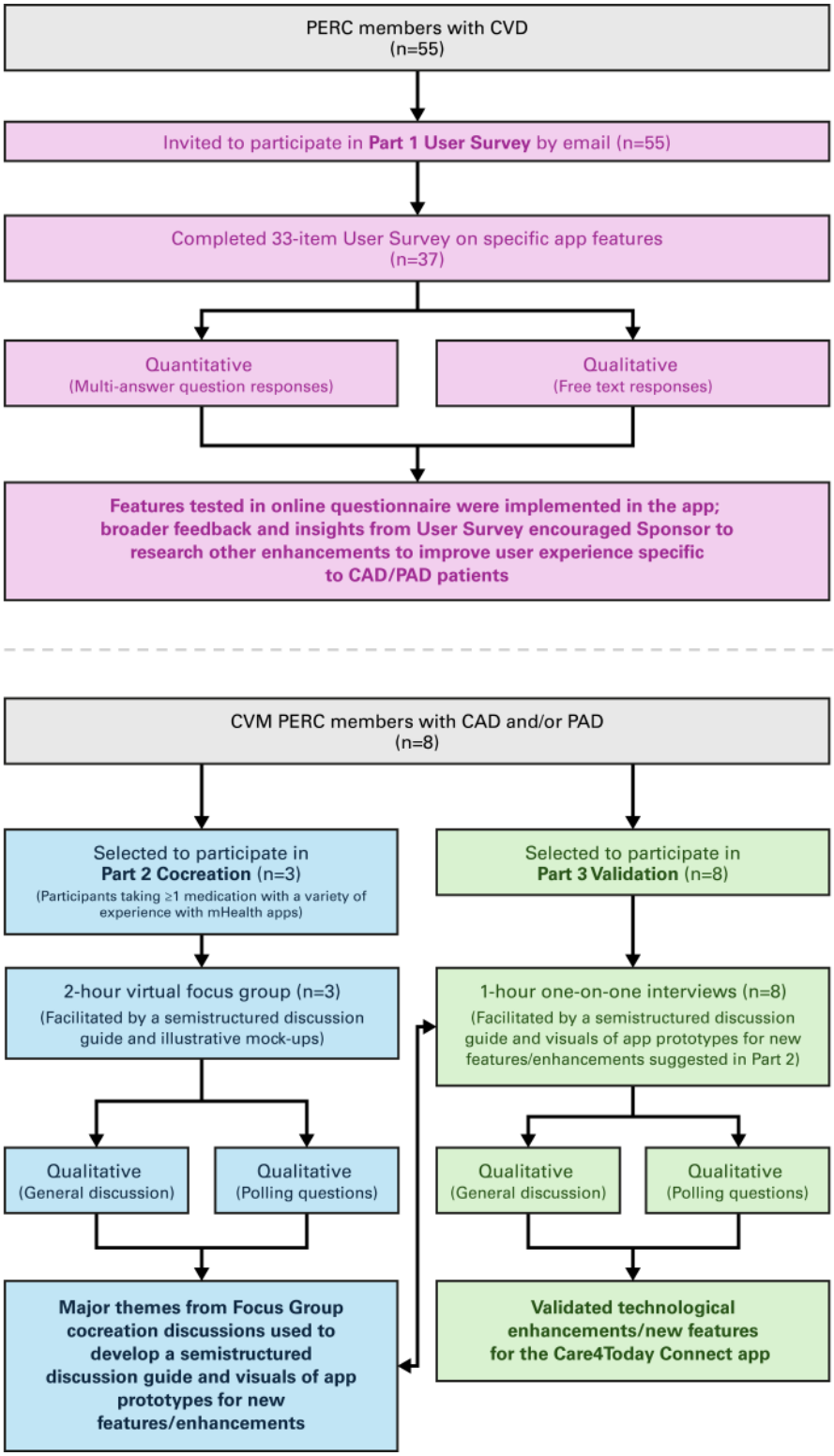


**Table 1.** Participant demographics.

| Characteristic | Part 1 (survey; n=37), n (%) | Part 2 (cocreation; n=3), n (%) | Part 3[a] (validation; n=8), n (%) |
|---|---|---|---|
| **Diagnosis[b]** | | | |
| CAD[c] and PAD[d] | —[e] | 1 (33) | 3 (38) |
| PAD alone | Not specified in the responses | 2 (67) | 5 (63) |
| CAD alone | — | 0 | 0 |
| **Sex** | | | |
| Male | 15 (41) | 1 (33) | 2 (25) |
| Female | 22 (59) | 2 (67) | 5 (63) |
| Nonbinary | 0 | — | 1 (13) |
| **Race** | | | |
| White | 10 (27) | 1 (33) | 4 (50) |
| Black | 23 (62) | 2 (67) | 3 (38) |
| Asian | 3 (8) | 0 | 1 (13) |
| Hispanic/Latino or Spanish in origin | 3 (8) | 0 | 0 |
| American Indian or Alaska Native | 1 (3) | 0 | 0 |
| Other | 1 (3) | — | 0 |
| **Age range (years)** | | | |
| 20-29 | 1 (3) | 0 | 0 |
| 30-39 | 1 (3) | 0 | 0 |
| 40-49 | 6 (16) | 1 (33) | 3 (38) |
| 50-59 | 6 (16) | 2 (67) | 1 (13) |
| 60-69 | 14 (38) | 0 | 3 (38) |
| 70-79 | 9 (24) | 0 | 1 (13) |
| **Highest education level** | | | |
| Less than high school | 1 (3) | 0 | 0 |
| High school | 2 (5) | 0 | 1 (13) |
| Some college | 4 (11) | 1 (33) | 1 (13) |
| Trade or technical school | 2 (5) | 1 (33) | 1 (13) |
| Bachelor’s degree | 13 (35) | 1 (33) | 4 (50) |
| Associate degree | 1 (3) | 0 | 0 |
| Graduate degree | 13 (35) | 0 | 1 (13) |
| Other | 1 (3) | — | 0 |
| **Region[f,g]** | | | |
| Urban | — | 1 (33) | 3 (38)[f] |
| Suburban | — | 2 (67) | 1 (13) |
| Rural | — | 0 | 2 (25) |
| **Comorbidities[b,f]** | | | |
| Diabetes | — | 2 (67) | 5 (63) |
| Any | — | 1 (33) | 6 (75) |
| **Medications/day** | | | |
| <7 | — | 0 | 1 (13) |
| 7-14 | — | 2 (67) | 5 (63) |

| Characteristic | Part 1 (survey; n=37), n (%) | Part 2 (cocreation; n=3), n (%) | Part 3[a] (validation; n=8), n (%) |
|---|---|---|---|
| ≥15 | — | 1 (33) | 2 (25) |

[a]3/8 participants from part 3 participated in part 2.

[b]Self-reported diagnosis.

[c]CAD: coronary artery disease.

[d]PAD: peripheral artery disease.

[e]Not applicable.

[f]Data were unavailable for 2 participants in parts 2 and 3.

[g]1 participant responded "urban" but indicated they had previously been "rural."

**Table 2.** Medication tracking and health experience reporting behavior (parts 2 and 3; n=8).

| Participant (parts 2 and 3) | Medication tracking behavior | Health experience reporting behavior |
|---|---|---|
| A (parts 2 and 3) | • No adherence medication tracking[a]<br>• Uses MyChart for tracking right dosing and frequency for medications<br>• Manual pill box used in the morning and afternoon/evening | • Uses reminders on continuous glucose monitor and compression boot app devices<br>• Uses a journal to record health experiences to be discussed in next health care provider visit |
| B (parts 2 and 3) | • No adherence medication tracking[a]<br>• Uses MyChart for tracking right dosing and frequency for medications<br>• Sets up a smartphone alarm twice daily for the morning and afternoon/evening | • No health experience tracking or reporting |
| C (parts 3) | • No current medication tracking[a]<br>• Used to track medications on an app | • No current health experience tracking or reporting<br>• Used to track blood pressure, glucose, bloating, and heart rate on an app, but found it too time-consuming |
| D (part 3) | • No adherence medication tracking[a]<br>• Places pills in a high visibility area | • No health experience tracking or reporting |
| E (part 3) | • Uses a weekly pill organizer for drugs for the morning and afternoon/evening | • No health experience tracking or reporting |
| F (part 3) | • Uses calendar app, alarms, and reminders to track medication | • Keeps track of health experience as part of morning routine<br>• Tracks blood pressure, glucose, time in range, weight, and pain on calendar app |
| G (parts 2 and 3) | • Uses retail pharmacy app for tracking medications list | • No health experience tracking or reporting |
| H (part 3) | • Uses phone alarms<br>• Manual pill box used in the morning and afternoon/evening | • Uses health app for tracking glucose (<30 min/d)<br>• No other health experience tracking or reporting |

[a]Digital or nondigital.

## Part 1: Internet Survey

### *Sample Characteristics*

In part 1, a total of 67% (37/55) of participants with cardiovascular disease completed the survey (Table 1). In total, 59% (22/37) participants were female, 59% (22/37) participants were White, 27% (10/37) participants were Black or African American, 62% (23/37) participants were aged ≥60 years, and 81% (30/37) participants had been educated beyond high school. Most (28/37, 76%) had been managing their health condition for >5 years. Overall, 78% (29/37) of survey respondents reported using mHealth apps at least once during the day to help manage their condition, with 35% (13/37) respondents reporting that they used mHealth apps somewhat or very often.

### *Understanding of Existing App Features*

When presented with the "Tutorial for New Users" feature (Multimedia Appendix 2), 70% (26/37) of respondents indicated they would continue to use the feature, rather than skip it, and expressed a high level of understanding at each step of the tutorial. Confidence in navigating to various tabs within the Care4Today Connect app was high and most (28/37, 76%) felt at least somewhat likely to use the app after the tutorial.

Respondents understood the concept of the "Earned Points" feature (Multimedia Appendix 2) and most (29/37, 78%) were confident in earning points when using the app but questioned the value of the points reward system. They considered the true value of the app to be in its ability to streamline the functionality of many apps they might be using into one.

> *Earning points may be motivation for using the app. However, the ability to condense what several apps do into 1 app for me would be a higher motivation. [It] would be nice to focus on that as a convenience and usability feature.* [Female participant, 60-69 years, cardiovascular and metabolic disease]

### App Use in Clinical Study

Respondents were asked to assume they had enrolled in a clinical trial that used the Care4Today Connect app and to consider what might drive them to use the app. Motivating factors included contributions to research (33/37, 89%), helping others (29/37, 78%), learning about health/disease (29/37, 78%), improving health (26/37, 70%), better disease management (25/37, 68%), and helping track medications (19/37, 51%). Potential drivers for not using the app included concerns around confidentiality/health data privacy and time obligation.

> *I would want control of when data is sent to my health care providers and who is authorized to receive that data.* [Male participant, 60-69 years, bladder cancer]
>
> *If it is a huge time obligation, or if it doesn't sync with my watch, or if it means that I still have to use multiple other apps that I already use on a daily basis...* [Female participant, 20-29 years, pulmonary hypertension]

Most participants thought the app would be useful in monitoring self-tracked health metrics, such as blood pressure or pain (29/37, 78%), health trends and progress (28/37, 76%), and lifestyle habits (27/37, 73%) (Table 3). Additionally, on a scale of 1 to 10 (where 1 was unlikely and 10 was very likely), most participants (28/37, 76%) selected a response of 7 or higher, indicating that they were likely to recommend the app to a friend or coworker.

**Table 3.** Self-track features of the Care4Today Connect app considered by participants as useful (part 1; n=37). Also, more than 1 item could be selected.

| Activity | Respondents, n |
|---|---|
| Tracking health metrics (eg, blood pressure, pain) | 29 |
| Monitoring my health trends and progress | 28 |
| Tracking lifestyle habits (daily routine, step count, mood, and sleep) | 27 |
| Learning new information about my health | 27 |
| Refreshing my knowledge on my health | 19 |
| Remembering when my medical appointments are scheduled | 18 |
| Remembering to take my medication as prescribed | 17 |
| Other | 4 |

## Part 2: Cocreation

### Sample Characteristics

Three participants from the CAD- or PAD-specific PERC were selected to participate in the virtual focus group in part 2, including 1 male and 2 female patients who were aged between 40 and 59 years, and all of whom were taking 7 or more medications (Table 1).

### Adding Medication Data for Tracking Features

Currently, adding medication data to the Care4Today Connect app involves manual input of multiple fields to create a customized experience for medication tracking (Multimedia Appendix 3). Illustrative mockups of potential features designed to enable auto-adding medication data to the app were shared with the 3 focus group participants (Multimedia Appendix 4).

Participants saw value in both options to auto-add medication data into the app. Adding medication data via a third-party insurance portal (option 1) was considered the most appealing and convenient solution for the initial setup, allowing a significant number of medications to be added at the same time. Adding medications with OCR technology (option 2) was not considered suitable for initial medication upload due to the associated time burden for patients with CAD, PAD, or both who are typically taking multiple medications; however, this feature was thought to be a better, more intuitive, simpler alternative to option 1 for subsequent additions and changes to medication lists.

> *For the initial setup, have it imported from your doctor's office…I would use both but initially I wouldn't want to take photos of 9 or 10 different bottles to set it up.* [Female participant, age 50-59 years, PAD]

Other suggestions for simplifying the process of adding medication data included integration with other medical apps, such as MyChart, and pharmacies. Of the 3 focus group participants, 2 already used MyChart for tracking the dosing and frequency of their medications, refills, setting and tracking appointments, contacting their HCPs, and reporting their health experiences (Table 2). Participants also suggested that connection to pharmacies to add medication data may be useful because pharmacy records are typically updated faster than electronic health record data.

### *Self-Reporting Health Experience Features*

Participants considered it a simple process to set up the Care4Today Connect app to track and self-report health experiences. While participants were sensitive to the burden of manual reporting of their health experiences, for 67% (2/3) of them this was outweighed by the perceived value of sharing data with their clinical team and having access to a record of their metrics. These 2 participants were willing to manually track their data for approximately 20 minutes/day, a time window corresponding to their disease management routines.

Reminder notifications on apps were considered critical for ongoing tracking, especially when set up to correspond with existing routines. Snooze and follow-up alarm functionalities were requested, rather than a single reminder. Participants expressed a strong preference for wearables or smart devices to overcome the burden of tracking.

> *Before I track any of these manually, I'd get one of those [smart] watches so it's tracking for me.* [Male participant, age 60-69 years, PAD and CAD]

Participants could also see the benefits of immediate feedback based on their self-reported health experiences, such as an automated notification to follow-up with their care team for high blood pressure. A strong preference for HCP-driven notifications and feedback was noted as participants expressed concerns and distrust over generic automated algorithm-generated notifications.

> *I want the doctors to look at my data and be able to intervene if [that is] something that's about to happen.* [Female participant, age 40-49 years, PAD]

> *No [I would not trust the generic notifications]. That type of stuff will have to come from the doctor. I don't know who's behind that information.* [Female participant, age 40-49 years, PAD]

### *Other Areas for Improvement*

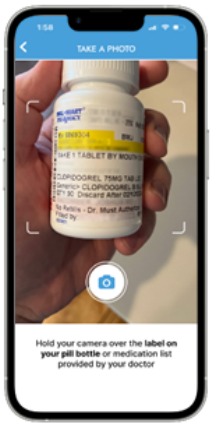
Other improvements to the Care4Today Connect app were discussed during the focus group. Areas of concern cited by participants included font sizes, number of fields, and amount of typing required to log into insurance portals to add medication data. One of the recommendations was a multiplatform/multidevice app with the ability to sign into a cloud portal via a laptop or tablet, which would have a larger screen and easier-to-use keyboard than a smartphone. A conversational AI interface was suggested, with language that makes reporting health experiences easier and more natural, intuitive, and engaging (eg, "Are your legs hurting today?" vs "Please report leg pain today").

> *We don't want to load data through the keyboard on [the] phone…most of us are down to 1 or 2 fingers using that. It's a very slow process as opposed to a regular keyboard even if we're still just using the same 2 fingers.* [Male participant, age 60-69 years, PAD and CAD]

## Part 3: Validation

In part 3, we sought to validate the additional enhancements proposed during part 2 (Figure 2) through 8 individual interviews with participants with CAD, PAD, or both, using a semistructured discussion guide and visuals of prototypes (Multimedia Appendix 6).

**Figure 2.** Care4Today Connect app: new features proposed for release 1.0 in CAD or PAD (part 2; n=3). CAD: coronary artery disease; HCP: health care provider; OCR: optical character recognition; PAD: peripheral artery disease; PRO: patient-reported outcome.

| | Medication Tracker set-up | | Health experience tracker | "FOR YOU" notification tab |
|---|---|---|---|---|
| **Feature** | Automated adding of medication data | Adding of medication via OCR technology | Inclusion of CAD/PAD health metrics to existing health experience trackers | Trends and insights into medication and health metrics |
| **Value** | Ease burden of loading medications and setting upmedication tracking and reduce manual input error | | Provide HCPs with PROs predictive of disease progression | Enhance engagement, motivation, and, potentially, patient outcomes |
| **Element** | Import patient medication list from a trusted third party (e.g., insurance portal) | Add or change medications from an image of medication bottle | Comprehensive new set of CAD/PAD health metrics | Personalized feedback on self-tracked data (trends/graphs and recommendations), plus appointment and medication refill reminders |
| **Prototype** |  |  |  |  |

### *Sample Characteristics*

Eight of the sponsor's cardiovascular PERC participated in part 3 of the research. Nearly two-thirds (5/8, 63%) of those interviewed were female, half (4/8, 50%) were White, most (7/8, 88%) had been educated beyond high school, and around two-thirds (5/8, 63%) were aged 50 years and older. Comorbidities were common, with 63% (5/8) reporting comorbid diabetes, and rates of polypharmacy were high, with 88% (7/8) of participants routinely taking ≥7 medications/day. Most participants used an app or a manual pill box as a reminder to take their medication, but only 50% (4/8) of individuals tracked their medications and fewer (3/8, 38%) tracked and reported their health experiences (Table 2).

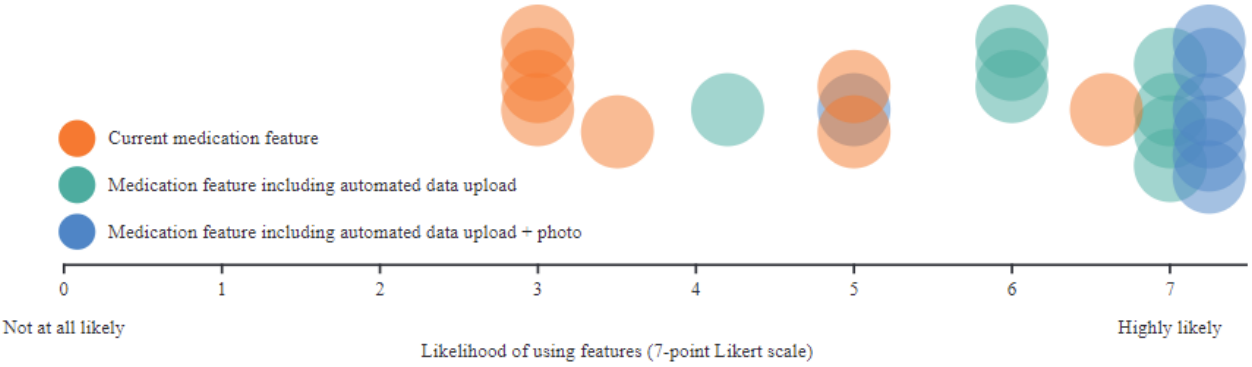
### *Adding Medication Data Feature*

In a study population with marked polypharmacy, participants found value in an app that helps them manage their medications together.

> *This would be very helpful because it would be all my medications and not just some.* [Female participant, age 50-59 years, PAD]

Both proposed medication data features (insurance portal and OCR technology) were well received. When rating the perceived value of each new feature, most participants reacted neutrally to the current manual method of adding medication data for tracking (Figure 3). Most respondents taking ≥7 medications/day (6/7, 85% of participants) considered automatically added medication data from a trusted third party (eg, insurance portal) to be a highly valuable feature. The remaining participant, who was taking 5 medications/day, preferred to upload medications to the app manually rather than via the 2 new features.

**Figure 3.** Perceived value of and willingness to use Care4Today medication upload features (part 3; n=8). Each circle represents a participant's response (3 per participant).




While the overall perception of using a third party to add medication data was positive, respondents flagged several potential barriers to its use. Two participants voiced concerns about medication accuracy due to delays in changes to medications on the provider portal. Two participants also worried about the accuracy of medication lists if insurance providers and pharmacies mix claims or if their medication records are not up to date. Another participant mentioned that their small insurer may not be connected to the app.

All respondents saw value in adding medication data via OCR technology for new medications or medication changes. Most found this approach to be preferable to manually typing on their smartphone, particularly due to dexterity issues caused by old age or disease. Only 1 participant felt taking an image of their medication bottle would be difficult, due to shaking hands.

All participants expressed the need to have both options included. Most (6/7, 85%) participants stated that the inclusion of these features increased the likelihood that they would use the Care4Today Connect app.

### *Self-Reporting Health Experiences Feature*

Participants reacted positively to a dedicated tool for tracking and sharing their CAD, PAD, or combined health experiences with their HCPs. Most expressed regret about having inaccurate discussions in their HCP visits due to gaps in their self-tracking of health metrics and experiences.

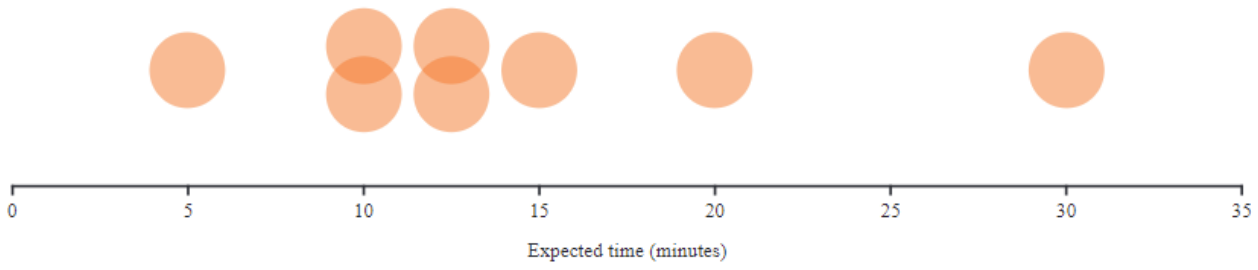
> *I want to start tracking my symptoms when and where they occur because my doctor does not believe me when I tell him.* [Female participant, 40-49, PAD]
>
> *I often forget what happened last week or last month, so I don't discuss my old symptoms with my doctor.* [Female participant, 60-69, CAD and PAD]

Presented with mock-ups of a conversational AI interface for tracking their health experiences (Multimedia Appendix 7), most participants indicated they preferred the more traditional route (ie, inputting data into fields or selecting options). The conversational interface was perceived to be impractical overall, and unsuitable for the reporting of health experience metrics.

> *I do not want to have a conversation, just log my data.* [Male participant, 70-79, CAD and PAD]
>
> *The text one feels less private.* [Female participant, 50-59, PAD]

Despite understanding the value of tracking and reporting their data to share with their HCPs, participants were hesitant to dedicate a significant amount of time to this. On average, they were willing to spend 10-15 minutes a day tracking their health metrics and experiences (4 or 5 metrics/day), including medications (Figure 4).

**Figure 4.** Expected time dedicated to reporting of health metrics and experiences, including symptom and medication reporting (part 3; n=8). Each circle represents 1 participant.




> *I'd say 10 or 15 minutes. I mean, that's pretty fair. We waste 10 or 15 minutes a day playing our little online games or something like that, so why not do something that could possibly benefit us, especially if the doctors are directly linked to the app?* [Female participant, age 40-49 years, PAD]

Certain health data and experiences were more likely to be tracked and reported than others, particularly those recommended by HCPs (Table 4). Participants with existing

health or medication routines were most likely to track and report their health experiences.

**Table 4.** Number of participants likely to self-track and report patient-reported outcomes (part 3; n=8).

| Health metric | Participants, n |
|---|---|
| Blood pressure | 6 |
| Glucose levels | 5 |
| Chest/leg pain or discomfort | 5 |
| Swollen feet and limbs, bloating | 2 |
| Sleep | 2 |
| Weight | 2 |
| Heart rate | 1 |
| Cramping | 1 |
| Shortness of breath | 1 |
| Palpitations | 1 |

### "For You" Section

A "For You" section in the app was considered essential to create an all-encompassing platform for managing CAD, PAD, or both. Personalized notifications were believed to be of value if they were validated by an HCP, rather than being an automated response from the app. These might include recommendations or actions for a particular health metric (eg, go for a walk), alerts to contact the care team (eg, call the nurse or schedule an appointment), and changes in medication.

> *I like it because you're directly communicating with your doctor instead of waiting a month in pain.* [Female participant, age 40-49 years, PAD]

> *I think it [rule-based notifications] would still be appreciated, but I think it would be deeply appreciated coming from the provider's practice.* [Female participant, age 60-69 years, PAD and CAD]

### Other Areas for Improvement

The user interface of the Care4Today Connect app was well received, particularly because of its simplicity, design, and intuitive workflow. There was an agreement with earlier feedback from the part 2 virtual focus group that the visual design could be improved by increasing button size, font size, and font contrast, and altering colors, to address accessibility and visibility challenges.

## Discussion

### Principal Results

This mixed methods research identified technological app enhancements to the Care4Today Connect, including improving the utility of the medication tracking as well as improving self-reported health data and experiences with relevant care team feedback, to optimize its ability to meet the specific requirements of patients with CAD, PAD, or both.

The existing Care4Today digital platform is continuously updated to enhance its utility. An initial internet survey of a broad group of cardiovascular participants, including those with CAD and PAD indicated a general understanding of key features, as well as opportunities for further enhancements. Based on this, we asked participants with CAD, PAD, or both for suggestions on improving the app. In a virtual focus group and individual interviews, participants told us they could see value in using technology to help add their medication data for tracking because it could reduce the user burden of having to manually add medication data. Participants also indicated that they would find self-reporting their health experiences valuable if the time obligation was not onerous. Further, respondents were interested in personalized in-app feedback from their care team based on their self-reported medication tracking and health experiences.

### Comparison With Prior Work

#### *Adding Medication Data Feature*

Polypharmacy is common in the CAD, PAD, or both populations, who typically comprise an older cohort with multiple comorbidities. In our focus group sample, 7/8 participants reported taking ≥7 medications/day, with 2 participants taking >15 medications/day. This is consistent with data from a claims-based study (n=148), in which 91% of patients with CAD were found to be taking ≥5 medications, with 74% taking ≥5 cardiovascular medications [10].

Polypharmacy is linked to both medication errors [11] and nonadherence [12]. mHealth apps provide a patient-centered means of targeting medication adherence [22]. Participants in our study stated that they would welcome multiple features on the Care4Today Connect app to allow for automated medication data to assist with medication tracking. Minimizing the reliance on manual input of data, by offering automated options, should reduce both the time burden associated with manual input and the potential for data entry errors. While older adults with CAD are proficient users of mobile apps and find them useful for medication adherence [23], our research highlights visibility and dexterity challenges as barrier to their use, particularly on a small screen. Automatic addition of medications using OCR technology has been shown to track medication adherence accurately [24], and optimization and flexibility of medication

data input are commonly requested by users of medication adherence apps [22].

### *Self-Reporting Health Experiences Feature*

A recent poll suggests that 2 in 5 US adults use mHealth apps, with at least half of them using the technology daily [25]. There is clear familiarity with this kind of technology among the general population and evidence of improvements in adherence and short- and long-term outcomes in people with CAD, PAD, or both who use mHealth apps [14-17]. Nevertheless, many of those in our study were either not currently tracking their medication and health experiences or were tracking these metrics through different channels or methods, such as pill boxes. In total, 78% of participants in part 1 said they used mHealth apps to help manage their disease, but only half the patients with CAD, PAD, or both in parts 2 and 3 reported routinely tracking their medications, with even fewer tracking and reporting their health experiences. Time constraints were identified as a barrier.

The ability to connect with their HCP or clinical team was positively received and participants were interested in additional notifications if they came with a personalized recommendation from their HCP. Immediate feedback on health metrics can enhance user engagement, motivation, and, potentially, patient outcomes by providing the user with a sense of progress. Indeed, a questionnaire-based survey of 180 patients with PAD concluded that information, monitoring, and feedback were the most relevant mHealth app components for this population [26].

## Strengths and Limitations

Patient feedback is essential for the optimization of the content and quality of digital health tools. The cocreation and validation approach used in our research ensured that participants with CAD, PAD, or both were involved in the co-design and refinement of potential enhancements to the Care4Today digital platform that would address their unique needs. Both quantitative and qualitative components ensured that valuable patient insights and rich context around their choices were captured to guide future app development. However, this was exploratory research and, as such, had several limitations. First, as with many studies of this nature, our focus group and interviews involved only a small number of participants with CAD or PAD, or both. Hence, our findings are not generalizable to t the broader population with CAD, PAD, or both. Second, while the study sample was ethnically and demographically diverse, participants had been invited to participate from existing PERC programs and, as such, self-selection bias resulted in a sample of participants that were more engaged and aware of their disease than the wider population of those with CAD or PAD, or both. This could potentially influence responses toward greater familiarity with mHealth apps. Third, CAD and PAD diagnoses were self-reported. There is a risk that self-reported diagnoses may differ from clinical diagnoses depending on the quality of patient–clinician communication, time since diagnosis, and the health literacy of the patient. Finally, employees of the Sponsor were present during virtual sessions. However, CorEvitas and ZS Associate researchers introduced themselves including first name, company affiliation, and research objectives. The facilitator's introduction informed participants of sponsor's presence but also included instruction that the aim was to gather participants' honest feedback and there were no wrong answers.

## Future Directions

Patients and HCPs are key stakeholders of any digital health tool, including the Care4Today Platform. Feedback from both groups is inherent to the digital platform's usability and adaptability across the health care system. While this article has focused on the patient, the Care4Today team has also engaged key opinion leaders in the Cardiovascular space, which include HCPs and professional organizations. There is an opportunity to take learnings from both engagements and explore a study where codevelopment and validation are conducted with both patients and HCPs.

## Conclusions

The Care4Today digital platform is focused on improving medication adherence, enabling self-reporting of health experiences providing patient education, enhancing connection with HCPs, and facilitating data and analytics learning across select disease areas. Our exploratory mixed methods research sought to identify how to improve the overall experience of patients with CAD or PAD, or both using Care4Today Connect. The goal was to understand patient insights and preferences on how they could add and self-report medication and health experience data. Key takeaways include recommendations to focus on enhancements that could reduce user burden through automation and technology, and foster HCP connection with personalized feedback. Incorporating new features that have been ideated and validated by patients, who are also end users, is crucial to the development and utility of digital apps. Through this research, the Care4Today team can prioritize the next iteration of the platform to optimize the experience for both patients and health care teams.

## Acknowledgments

Sponsorship for this study as well as all publication charges were funded by Janssen Scientific Affairs, LLC, a Johnson & Johnson company. The authors thank the patients who participated in the Patient Engagement Research Council (PERC) activities for their engagement and insightful feedback. Melanie Jones, BSc, and Kelsey Hodge-Hanson, PhD, of Twist Medical provided medical writing and editorial assistance, funded by Janssen Scientific Affairs, LLC, a Johnson & Johnson company. Gunter Scherk and Brad Marcum supported data analysis.

## Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## Authors' Contributions

SJ, GG, AG, and ED were involved in the study design; AG and AED carried out data collection; and AG performed data analysis/interpretation. All authors reviewed and critically revised the manuscript and approved the final version. All authors agree to be accountable for the accuracy and integrity of the work.

## Conflicts of Interest

SJ, BW, AH, SK, MT, and GG disclose that they are employees of Janssen Scientific Affairs. BW is an employee of Johnson & Johnson Technology Services. AG is an employee of ZS Associates, and AED is an employee of CorEvitas, LLC. Both ZS Associates and CorEvitas derive profits from interactions with pharmaceutical sponsors. CS and GP are members of Janssen's Patient Engagement Research Council who have been diagnosed with coronary artery disease or peripheral artery disease, or both, and were compensated financially for their time.

Multimedia Appendix 1
Janssen's Cardiovascular Metabolic Patient Engagement Research Council: eligibility criteria.
[PDF File (Adobe PDF File), 70 KB - cardio_v9i1e56053_app1.pdf ]

Multimedia Appendix 2
Part 1 internet survey.
[PDF File (Adobe PDF File), 254 KB - cardio_v9i1e56053_app2.pdf ]

Multimedia Appendix 3
Screenshots of the existing feature for adding medication data manually on the Care4Today Connect app.
[PDF File (Adobe PDF File), 26 KB - cardio_v9i1e56053_app3.pdf ]

Multimedia Appendix 4
Illustrative mock-ups presented to facilitate discussion during the part 2 virtual focus group: (A) features for adding medication data and (B) features for health experience reporting.
[PDF File (Adobe PDF File), 157 KB - cardio_v9i1e56053_app4.pdf ]

Multimedia Appendix 5
Prototypes for the Care4Today Connect app presented during the part 3 interviews: (A) automatic adding of medication data concept testing and (B) tracking and sharing of health experiences value proposition testing.
[PDF File (Adobe PDF File), 126 KB - cardio_v9i1e56053_app5.pdf ]

Multimedia Appendix 6
Semistructured discussion agenda for individual validation interviews (part 3).
[PDF File (Adobe PDF File), 65 KB - cardio_v9i1e56053_app6.pdf ]

Multimedia Appendix 7
Example of a conversational user interface on a mobile health app.
[PDF File (Adobe PDF File), 92 KB - cardio_v9i1e56053_app7.pdf ]

## References

1. Tabi K, Randhawa AS, Choi F, Mithani Z, Albers F, Schnieder M, et al. Mobile apps for medication management: review and analysis. JMIR Mhealth Uhealth 2019;7(9):e13608 [FREE Full text] [doi: 10.2196/13608] [Medline: 31512580]
2. Al-Arkee S, Mason J, Lane DA, Fabritz L, Chua W, Haque MS, et al. Mobile apps to improve medication adherence in cardiovascular disease: systematic review and meta-analysis. J Med Internet Res 2021;23(5):e24190 [FREE Full text] [doi: 10.2196/24190] [Medline: 34032583]
3. What is coronary heart disease? National Heart and Lung Blood Institute. 2022. URL: https://www.nhlbi.nih.gov/health/coronary-heart-disease [accessed 2023-09-13]

4. Bauersachs R, Zeymer U, Brière JB, Marre C, Bowrin K, Huelsebeck M. Burden of coronary artery disease and peripheral artery disease: a literature review. Cardiovasc Ther 2019;2019:8295054 [FREE Full text] [doi: 10.1155/2019/8295054] [Medline: 32099582]
5. Poredos P, Jug B. The prevalence of peripheral arterial disease in high risk subjects and coronary or cerebrovascular patients. Angiology 2007;58(3):309-315. [doi: 10.1177/0003319707302494] [Medline: 17626985]
6. Brown JC, Gerhardt TE, Kwon E. Risk factors for coronary artery disease. In: StatPearls. Treasure Island, FL: StatPearls Publishing; 2023.
7. Ho PM, Magid DJ, Shetterly SM, Olson KL, Maddox TM, Peterson PN, et al. Medication nonadherence is associated with a broad range of adverse outcomes in patients with coronary artery disease. Am Heart J 2008;155(4):772-779. [doi: 10.1016/j.ahj.2007.12.011] [Medline: 18371492]
8. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? a systematic review of definitions. BMC Geriatr 2017;17(1):230 [FREE Full text] [doi: 10.1186/s12877-017-0621-2] [Medline: 29017448]
9. Wang X, Liu K, Shirai K, Tang C, Hu Y, Wang Y, et al. Prevalence and trends of polypharmacy in U.S. adults, 1999-2018. Glob Health Res Policy 2023;8(1):25 [FREE Full text] [doi: 10.1186/s41256-023-00311-4] [Medline: 37434230]
10. Tinoco MS, Groia-Veloso RCDS, Santos JNDD, Cruzeiro MGM, Dias BM, Reis AMM. Medication regimen complexity of coronary artery disease patients. Einstein (Sao Paulo) 2021;19:eAO5565 [FREE Full text] [doi: 10.31744/einstein_journal/2021AO5565] [Medline: 33729285]
11. Koper D, Kamenski G, Flamm M, Böhmdorfer B, Sönnichsen A. Frequency of medication errors in primary care patients with polypharmacy. Fam Pract 2013;30(3):313-319. [doi: 10.1093/fampra/cms070] [Medline: 23132894]
12. Pasina L, Brucato AL, Falcone C, Cucchi E, Bresciani A, Sottocorno M, et al. Medication non-adherence among elderly patients newly discharged and receiving polypharmacy. Drugs Aging 2014;31(4):283-289. [doi: 10.1007/s40266-014-0163-7] [Medline: 24604085]
13. Eckardt I, Buschhaus C, Nickenig G, Jansen F. Smartphone-guided secondary prevention for patients with coronary artery disease. J Rehabil Assist Technol Eng 2021;8:2055668321996572 [FREE Full text] [doi: 10.1177/2055668321996572] [Medline: 33796334]
14. Cruz-Cobo C, Bernal-Jiménez M, Vázquez-García R, Santi-Cano MJ. Effectiveness of mHealth interventions in the control of lifestyle and cardiovascular risk factors in patients after a coronary event: systematic review and meta-analysis. JMIR Mhealth Uhealth 2022;10(12):e39593 [FREE Full text] [doi: 10.2196/39593] [Medline: 36459396]
15. Kaihara T, Intan-Goey V, Scherrenberg M, Falter M, Frederix I, Dendale P. Impact of activity trackers on secondary prevention in patients with coronary artery disease: a systematic review and meta-analysis. Eur J Prev Cardiol 2022;29(7):1047-1056. [doi: 10.1093/eurjpc/zwab146] [Medline: 34472613]
16. Li Y, Gong Y, Zheng B, Fan F, Yi T, Zheng Y, et al. Effects on adherence to a mobile app-based self-management digital therapeutics among patients with coronary heart disease: pilot randomized controlled trial. JMIR Mhealth Uhealth 2022;10(2):e32251 [FREE Full text] [doi: 10.2196/32251] [Medline: 34906924]
17. Paldán K, Steinmetz M, Simanovski J, Rammos C, Ullrich G, Jánosi RA, et al. Supervised exercise therapy using mobile health technology in patients with peripheral arterial disease: pilot randomized controlled trial. JMIR Mhealth Uhealth 2021;9(8):e24214 [FREE Full text] [doi: 10.2196/24214] [Medline: 34398800]
18. Care4Today Connect. Care4Today. URL: https://www.care4today.com/cvd [accessed 2023-11-30]
19. Janssen Patient Engagement Research Councils. Janssen. URL: https://www.janssenpatientcouncils.com/index.html [accessed 2023-09-04]
20. Chakravarty SD, Abell J, Leone-Perkins M, Orbai A. A novel qualitative study assessing patient-reported outcome measures among people living with psoriatic arthritis or ankylosing spondylitis. Rheumatol Ther 2021;8(1):609-620 [FREE Full text] [doi: 10.1007/s40744-021-00289-w] [Medline: 33616863]
21. Kohli H, Agarwal J, Kumar M. An improved method for text detection using adam optimization algorithm. Glob Transitions Proc 2022;3(1):230-234 [FREE Full text] [doi: 10.1016/j.gltp.2022.03.028]
22. Park JYE, Li J, Howren A, Tsao NW, de Vera M. Mobile phone apps targeting medication adherence: quality assessment and content analysis of user reviews. JMIR Mhealth Uhealth 2019;7(1):e11919 [FREE Full text] [doi: 10.2196/11919] [Medline: 30702435]
23. Park LG, Ng F, K Shim J, Elnaggar A, Villero O. Perceptions and experiences of using mobile technology for medication adherence among older adults with coronary heart disease: a qualitative study. Digit Health 2020;6:2055207620926844 [FREE Full text] [doi: 10.1177/2055207620926844] [Medline: 32489672]
24. Sarzynski E, Decker B, Thul A, Weismantel D, Melaragni R, Cholakis E, et al. Beta testing a novel smartphone application to improve medication adherence. Telemed J E Health 2017;23(4):339-348. [doi: 10.1089/tmj.2016.0100] [Medline: 27564971]
25. National tracking poll #2301120. Morning Consult. 2023. URL: https://pro-assets.morningconsult.com/wp-uploads/2023/02/16130638/2301120_crosstabs_MC_HEALTH_DIGITAL_HEALTH_DEVICES_Adults.pdf [accessed 2023-09-18]
26. Lortz J, Simanovski J, Kuether T, Kreitschmann-Andermahr I, Ullrich G, Steinmetz M, et al. Needs and requirements in the designing of mobile interventions for patients with peripheral arterial disease: questionnaire study. JMIR Form Res 2020;4(8):e15669 [FREE Full text] [doi: 10.2196/15669] [Medline: 32663154]

## Abbreviations

**CAD:** coronary artery disease
**HCP:** health care provider
**HIPAA:** Health Insurance Portability and Accountability Act
**IRB:** institutional review board
**mHealth:** mobile health
**OCR:** optical character recognition
**PAD:** peripheral artery disease
**PERC:** Patient Engagement Research Council

*Edited by N Cahill; submitted 19.01.24; peer-reviewed by T Kaihara, G Goodman, A Keogh; comments to author 02.05.24; revised version received 03.07.24; accepted 24.10.24; published 17.03.25.*

*Please cite as:*
*Juan S, Harxhi A, Kaul S, Woods B, Tran M, Geonnotti G, Gupta A, Dean E, Saunders CE, Payne G*
*Optimization of the Care4Today Digital Health Platform to Enhance Self-Reporting of Medication Adherence and Health Experiences in Patients With Coronary or Peripheral Artery Disease: Mixed Methods Study*
*JMIR Cardio 2025;9:e56053*
*URL: https://cardio.jmir.org/2025/1/e56053*
*doi:10.2196/56053*
*PMID:*

# Patient and Clinician Perspectives on Alert-Based Remote Monitoring–First Care for Cardiovascular Implantable Electronic Devices: Semistructured Interview Study Within the Veterans Health Administration

Allison Kratka[1,2], MD; Thomas L Rotering[1,2], MPH; Scott Munson[2], BS; Merritt H Raitt[3,4], MD; Mary A Whooley[1,2], MD; Sanket S Dhruva[1,2], MD, MHS

[1]University of California, San Francisco School of Medicine, San Francisco, CA, United States

[2]Department of Medicine, San Francisco VA Health Care System, 4150 Clement St, San Francisco, CA, United States

[3]Section of Cardiology, Veterans Affairs Portland Health Care System, Portland, OR, United States

[4]Knight Cardiovascular Institute, Oregon Health & Science University, Portland, OR, United States

**Corresponding Author:**
Sanket S Dhruva, MD, MHS
University of California, San Francisco School of Medicine, San Francisco, CA, United States

## Abstract

**Background:** Patients with cardiovascular implantable electronic devices (CIEDs) typically attend in-person CIED clinic visits at least annually, paired with remote monitoring (RM). As the CIED data available through in-person CIED clinic visits and RM are nearly identical, the 2023 Heart Rhythm Society expert consensus statement introduced "alert-based RM," an RM-first approach where patients with CIEDs that are consistently and continuously connected to RM, in the absence of recent alerts and other cardiac comorbidities, could attend in-person CIED clinic visits every 24 months or ultimately only as clinically prompted by actionable events identified on RM. However, there is no published information about patient and clinician perspectives on barriers and facilitators to such an RM-first care model.

**Objective:** We aimed to understand patient and clinician perspectives about an RM-first care model for CIED care.

**Methods:** We interviewed 40 rural veteran patients who were experienced with RM with CIEDs and 22 CIED clinicians who were experienced in using RM regarding barriers and facilitators to an RM-first care model. We conducted a reflexive thematic analysis of interviews. Two authors familiarized themselves with the dataset and generated separate codebooks based on the interview guides and inductively coded notes. These 2 authors met and reviewed each other's codes, sought additional author input, and resolved differences before 1 author coded the remaining interviews and developed candidate themes. These themes were refined, named, and supported with quotations.

**Results:** Patients expressed interest in an RM-first approach, to reduce the burden of long travel times, sometimes in inclement weather, and to enable clinicians to provide care for other patients. However, many preferred routine in-person visits; reasons included a skepticism of the capabilities of RM, a sense that in-person visits provided superior care, and enjoyment of in-person patient-clinician relationships. Clinicians were interested in RM-first care, especially for stable, RM-adherent patients who were not device-dependent. Clinicians most frequently cited the benefit of reducing patient travel burden as well as optimizing clinic space and time to focus on other care such as reviewing routine RM transmissions, but also noted barriers including lack of in-person assessment, patient-perceived diminution of the patient-clinician relationship, possible loss to follow-up, and technological difficulties. Clinicians felt that an RM-first care model should be evaluated for success based on patient satisfaction and assessment of timely addressing of rhythm issues to prevent adverse outcomes. Most clinicians believed that RM-first care represented the future of CIED care.

**Conclusions:** Both patients and CIED clinicians interviewed who were experienced in using RM were open to an RM-first care model that reduces in-person visits but reported some barriers to solely relying on RM and possible diminution of the patient-clinician relationship. Implementation of new RM recommendations will require attention to these perceptions and prioritization of patient-centered approaches.

***(JMIR Cardio 2025;9:e66215)*** doi:10.2196/66215

**KEYWORDS**

cardiovascular implantable electronic device; CIED; remote monitoring; RM; alert-based monitoring; remote monitoring–first care; patient perspectives; clinician perspectives; veteran; pacemaker; implantable cardioverter-defibrillator; mobile phone

## Introduction

Remote monitoring (RM) is the standard of care for patients with cardiovascular implantable electronic devices (CIED; pacemaker or implantable cardioverter-defibrillator [ICD]) [1,2]. RM involves sending CIED data from a patient's residence via a transmitter or smartphone app. Routine transmissions are usually sent every 90 days and can also be patient- or alert-initiated. RM is a Class 1, Level of Evidence A, professional society recommendation because of its many clinical outcome benefits [1,2]. These include reduced mortality [3-5], fewer hospitalizations [3,6], fewer inappropriate ICD shocks [7], as well as high patient satisfaction [8].

In addition to RM, CIEDs can also be checked in person; traditionally, patients attend routine in-person clinic visits at least annually [1]. However, because nearly all of the same CIED-related data can be obtained via RM, an alternative would be to end in-person visits completely if patients were consistently and continuously connected to RM, with in-person evaluations only when needed for clinically actionable reasons, such as CIED reprogramming [2].

The 2023 Heart Rhythm Society (HRS) expert consensus statement on practical management of the remote device clinic introduced such a novel care model, "alert-based RM," in which patients with CIEDs that are consistently and continuously connected to RM, in the absence of recent alerts or other cardiac comorbidity, could attend in-person CIED clinic visits every 24 months (class 2a recommendation) [2]. This statement is supported by multiple randomized, controlled trials that have demonstrated no difference in cardiovascular events [2,9-11] while reducing in-person visits, loss to follow-up, staff workload, and costs of care [9-11].

Additionally, the professional society expert consensus discussed the possibility of ending all routine in-person visits, given that these visits may be "low-value" because most conclude that the CIED is working properly [2]. In-person visits would occur only as clinically prompted by actionable events identified on RM. Such an RM-first care model, where patients have routine in-person visits every 2 years, or even only as needed, if they remain consistently and continuously connected could be especially helpful for the Veterans Health Administration (VHA) patient population, because approximately 40% of veterans with CIEDs who participate in RM live in a rural area [12] (defined as a land area outside of a census tract with ≥30% of the population residing in an urbanized area as defined by the Census Bureau) [13] and often have long travel times to clinic visits.

Despite these potential advantages and the HRS recommendation supported by multiple randomized controlled trials, patient and clinician perspectives on this new care model have not been studied. To understand barriers and facilitators to implementation, we conducted a mixed methods evaluation to explore the perspectives of device clinicians and veterans with CIEDs on an RM-first care model.

## Methods

### Interview Guide and Survey Development

One semistructured interview guide for veteran patients and one for clinicians (Multimedia Appendix 1) was developed by the investigator team using the updated Consolidated Framework for Implementation Research [14]. The veteran interview guide was developed based on a prior veteran survey about RM [15] and revised with input from the Rural Colorado Veteran Research Engagement board. The clinician interview guide was developed through an iterative process with input solicited from practicing VHA cardiologists and the incorporation of concepts from new HRS recommendations [2].

Both interview guides sought to understand barriers and facilitators to an "RM-first strategy," defined as in-person CIED clinic visits only if clinically prompted among patients engaged in RM. Patients were informed that similar data were obtained through RM as in-person visits; they may need in-person visits for abnormalities identified on remote transmissions; they could still contact their device clinic; and their other visits, such as with primary care, would continue. Patients were asked about the travel burden to VHA, how their care may have changed during the COVID-19 pandemic, and any concerns about reducing routine in-person CIED clinic visits. Device clinicians were asked about the benefits and barriers to this new care model, and how this may change their practice flow. A 23-item Qualtrics survey was also administered to gather professional and demographic data as well as preinterview information about clinician impressions of RM-first care (Multimedia Appendix 1). Specifically, this survey asked clinicians how often they conducted routine evaluations for patients with CIEDs, stratified by adherent and nonadherent patients, and what clinicians did when patients did not want to schedule routine in-person CIED checks or missed an in-person CIED check. This survey also asked clinicians about the anticipated benefits and concerns of an RM-first strategy, how effective that it would be concerning cardiovascular outcomes, and if such a strategy would help their clinic.

Of note, partway through the clinician interview process, the draft 2023 HRS expert consensus was released [2], introducing an "alert-based care" model, similar to RM-first care. Therefore, the interview guide was then adapted to solicit feedback about this recommendation. For the veteran interviews, a question was added about the veteran's view of the new recommendations.

This was a quality improvement project conducted in partnership with the VHA Measurement Science Quality Enhancement Research Initiative and the VHA National Cardiac Device Surveillance Program.

### Study Population and Contact Process

Veterans were eligible for interview inclusion if they had a CIED, were completely adherent to RM in the past 400 days (which means that they had sent a remote transmission covering this timeframe), [12] and lived in a rural area. Introductory letters were sent to 100 randomly selected veterans meeting these criteria (since these participants did not know the project team), 91 of whom were then contacted at least once via a telephone connection to Microsoft Teams. The letter described the study background and objectives as well as topics that would be covered by a named VHA staff member (SM). Up to 3 contact attempts were made, with a message left for each unanswered attempt.

A purposive sample of VHA CIED clinic-focused clinicians who had been interviewed for a prior project about best practices to support RM adherence were contacted for interview [16]. An introductory email described this study's background, objectives, and potential changes that may result from findings as well as information about the project team and funding source. Snowball sampling was then used, asking these clinicians to recommend colleagues at their device clinic. Finally, purposive sampling was used to contact clinicians caring for a high proportion of veterans living in rural areas to more adequately represent rural clinician perspectives.

### Interview Process

Informed consent was obtained before recording all interviews, which were conducted on and recorded using Microsoft Teams. Between November 2022 and February 2023, a total of 40 veterans were interviewed by coauthor SM (BS, male, qualitative researcher), with each of these 40 individual interviews lasting 5 - 15 minutes in length and some attended by coauthors TLR (MPH, male, public health researcher) and SSD (MD, MHS, male, cardiologist). Between November 2022 and February 2023, a total of 22 clinician interviews between 30 - 60 minutes were conducted by TLR, with some attended by SSD. Field notes were taken during both sets of interviews to summarize key points and supplemented with transcribed interview recordings to ensure accuracy. There were no repeat interviews.

### Qualitative Data Analysis

Reflexive thematic analysis [17,18] of interview field notes and transcripts was used to elucidate veteran and clinician views about RM-first care.

First, authors AK (MD, female, cardiology fellow) and TLR familiarized themselves with the dataset by reading the field notes and transcripts, making notes about the overall findings within both sets of interviews (veteran and clinician) and reflecting on their experiences in the direct care of patients with CIEDs (AK) and research and quality improvement efforts for care of patients with CIEDs (TLR). Next, the authors generated separate codebooks based on the domains of the distinct interview guides. For veteran interviews, AK and TLR independently coded 6 distinct interview notes, which involved generating additional codes identified inductively, for the goal of reflexivity. These 2 authors then met and reviewed each other's codes, sought SSD's input, and resolved any differences by consensus, creating 1 final codebook. AK then coded the remaining interviews and developed candidate themes, supporting each theme based on coded data and direct quotations. AK's candidate themes were intentionally broad. TLR and SSD reviewed these themes with AK against the coded data, leading to refining and then naming these themes. Finally, AK wrote the analytic narrative and supported these themes with quotations directly from the veteran interviews to describe veteran perspectives. Coauthor SSD provided iterative feedback on several versions of the analytic narrative to improve clarity and increase confirmability.

For clinician interviews, AK and TLR first independently coded 3 distinct interview notes, which involved generating additional codes identified inductively. These 2 authors then reviewed each other's codes and resolved any differences by consensus. AK then coded the remaining interviews. The authors used the same process as described above for thematic generation, refinement, and naming. AK wrote the analytic narrative, which is presented in the Results section of this paper, and supported these themes with quotations directly from the interviews. We conducted both clinician and patient interviews until reaching thematic saturation on two criteria, (1) no new concepts were identified in iterative analysis interviews (code frequency counts) and (2) there was consistent repetition among interviewee responses without any new information being added to existing codes (code meaning) [19,20]. The number of interviews that we conducted with both our population of veterans and Veterans Affairs (VA) clinicians exceeded the number (n=17) found in recent empiric studies [20].

Atlas.ti 23 (ATLAS.ti Scientific Software Development GmbH), a qualitative analysis software, was used to organize and apply analytic codes.

### Ethical Considerations

This work was conducted as a quality improvement project and not human subjects research. Per the Department of Veterans Affairs Office of Research & Development Program Guide: 1200.21, "VHA (Veterans Health Administration) Operations Activities That May Constitute Research," data were collected as part of a quality improvement study to assess and improve the quality of RM care for veterans with CIEDs and did not require institutional review board approval. Veteran and clinician participants were informed at study enrollment that responses would be anonymized, and verbal consent to recording was acquired before each interview. No compensation was provided. Study data were deidentified and stored in a secure, encrypted VA database.

## Results

### Veteran Interviews

#### *Overview*

Among the 100 veterans who were initially mailed a letter to request participation, for patient sex, 97 (97%) were male and 3 (3%) were female; for patient race, 2 (2%) were American Indian or Alaska Native, 7 (7%) were Black or African-American, 3 (3%) were Native Hawaiian or other

Pacific Islander, 81 (81%) were White, 5 (5%) declined to answer, and 1 (1%) was unknown; and for patient ethnicity, 1 (1%) was Hispanic or Latino, 96 (96%) were not Hispanic or Latino, 1 (1%) declined to answer, and 2 (2%) were unknown. Of 45 veterans contacted, 40 agreed to an interview (5 declined; Figure 1). The mean patient age was 77.6 (SD 8.9) years and all 40 were male (Table 1).

For their current care, most patients reported attending routine in-person visits to have their CIED checked (Table 1), usually every 6 - 12 (range 2 - 12) months. Many patients bundled other in-person VHA visits for convenience. Most patients did not think the COVID-19 pandemic had significantly changed their current CIED care.

When asked about an RM-first care model, 4 veterans preferred RM-first, 16 were amenable, 2 had no preference, and 18 did not want it. When asked what feedback they would prefer in an RM-first care model, few veterans wanted to know only when there was a problem, whereas more wanted feedback regarding successful or normal transmissions. The themes of barriers and facilitators to RM-first care described by veterans are in Table 2.

**Figure 1.** Flow diagram for veteran contact. CIED: cardiovascular implantable electronic device; RM: remote monitoring.

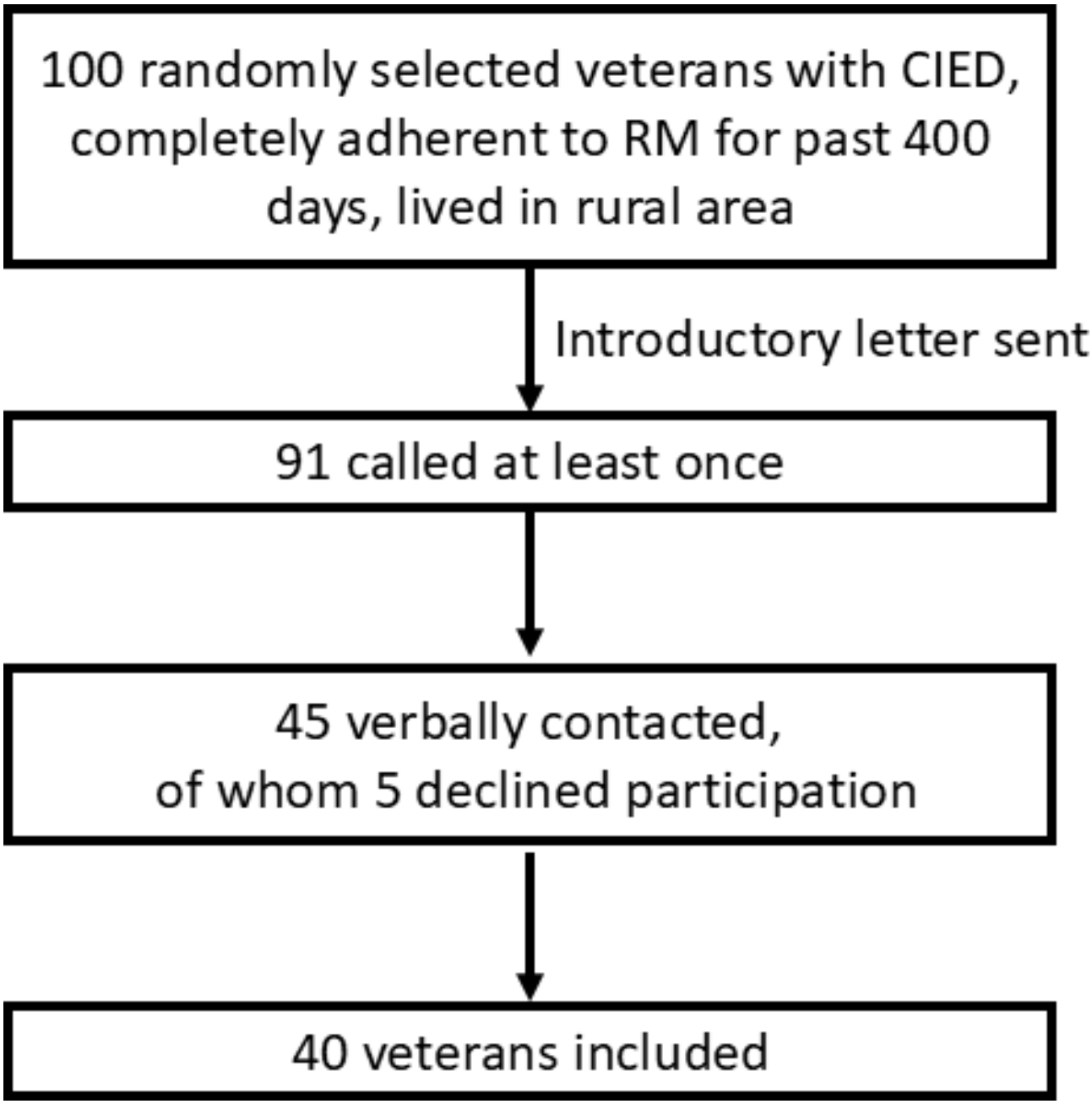


**Table .** Characteristics of veterans interviewed (n=40).

| | | Veterans interviewed |
|---|---|---|
| Age (years), mean (SD) | | 77.6 (8.9) |
| Gender, n (%) | | |
| | Male | 40 (100) |
| | Female | 0 (0) |
| Race, n (%) | | |
| | American Indian or Alaska Native | 1 (2) |
| | Black or African American | 2 (5) |
| | Native Hawaiian or other Pacific Islander | 1 (2) |
| | White | 35 (88) |
| | Declined to answer | 1 (2) |
| Ethnicity, n (%) | | |
| | Hispanic or Latino | 0 |
| | Not Hispanic or Latino | 39 (98) |
| | Unknown | 1 (2) |
| Type of device, n (%) | | |
| | Implantable cardioverter-defibrillator | 18 (45) |
| | Pacemaker | 22 (55) |
| | Wireless-capable device[a] | 34 (85) |
| Attended an in-person device clinic visit in the past year, n (%) | | |
| | Yes | 23 (58) |
| | No | 17 (43) |
| Attended a telephone device clinic visit in the past year, n (%) | | |
| | Yes | 28 (70) |
| | No | 12 (30) |
| Attended a VA[b] Video Connect device clinic visit in the past year, n (%) | | |
| | Yes | 3 (8) |
| | No | 37 (93) |
| Travel time to the VA (time for 1-way trip), n (%) | | |
| | Less than 1 h | 17 (42) |
| | 1 - 2 h | 15 (38) |
| | 2 - 3 h | 6 (15) |
| | More than 4 h | 2 (5) |
| Patient-reported frequency of in-person device clinic visits, n (%) | | |
| | Every 2 - 3 weeks | 1 (2) |
| | Every 2 months | 2 (5) |
| | Every 3 - 4 months | 6 (15) |
| | Every 6 months | 13 (32) |
| | >6 months and <1 year | 3 (8) |
| | Every year | 13 (32) |
| | Not available | 2 (5) |

[a]For context only, the 6 devices that were not wireless-capable were all pacemakers.

[b]VA: Veterans Affairs.

**Table .** Themes of barriers and facilitators to remote monitoring-first care.

| Barriers | Facilitators |
|---|---|
| Veterans | |
| Importance of in-person care | Travel burden |
| Concerns about the adequacy of RM[a] technology for care | Weather-related concerns |
| Loss of clinician-patient relationship | Comfort with technology |
| N/A[b] | Reducing the burden on the VHA[c] device clinic |
| Clinicians | |
| Benefits of routine in-person assessment | Reduced veteran travel burden |
| Reducing veteran contact with VHA | Optimization of clinic space and clinic staff time |
| Clinic operations-related changes | More time to review routine transmissions and improve RM adherence |
| Technology and technological difficulties for veterans and clinicians | No concern about relative value units |

[a]RM: remote monitoring.

[b]N/A: not applicable.

[c]VHA: Veterans Health Administration.

### *Barriers to RM-First Care*

#### Importance of In-Person Care

Many patients who were not amenable to RM-first care believed that in-person evaluations provided more valuable information and essential care that could not be obtained another way. As one veteran stated,

> *In person… they take a lot of recordings and stuff when they check the defibrillator… I think that it is [more accurate].*

#### Concerns About Adequacy of RM Technology for Care

Many veterans expressed concerns about the adequacy of RM technology for care. For some, this was based on a lack of comfort and sometimes a lack of confidence in RM technology or a belief that they needed more care because they had serious cardiac conditions.

> *[Remote monitoring] is a good idea if we can understand what to do with the electronics… That is a little difficult for us.*

Some of these concerns may stem from an expressed lack of information about the capabilities of RM, what parameters are obtained from RM, and what clinicians do with that information.

> *I'm not sure how they can check my [device] with the online system that I have…I don't see how they would do it virtually, because they usually have to put a wand over the pacemaker to check its function.*

#### Loss of Clinician-Patient Relationship

A few patients noted that the loss of their relationship with their clinician would be a barrier to an RM-first care model.

> *I actually look forward to the patient to doctor type meetings… there's something to be said about personal visits.*

### Benefits of RM-First Care

### *Travel Burden and Weather-Related Concerns*

Many veterans noted less time and cost burden would be required for travel to their VHA facility. For a few patients, this was related to poor mobility.

> *It saves me 100 miles of driving, and if we can accomplish the same thing, I think that would be a lot better.*
>
> *I don't have to spend an hour on the highway and save on gas too.*

For some veterans, this travel burden was sometimes due to weather-related issues.

> *It's a little bit because of the snow and weather here in Montana, and the pass that I have to go over to get to the VA.*

### *Comfort With Technology*

Several veterans did not have concerns regarding reduced quality of care with forgoing routine in-person visits and were comfortable with the quality of RM. As one veteran stated,

> *The technology is going to continue to improve. And those monitors are just going to get better and better. So that really eliminates the need to go inside and talk to the technician… If I don't have to [go to face-to-face visits], you're not exposing yourself to other patients being sick and all that.*

Some veterans felt reassured that RM would adequately monitor their device.

> *I think it would be alright as long as I know they're checking my machine and make sure it's up running.*

### *Reducing Burden on the Clinic*

Some patients mentioned that this new model of care would reduce the burden on their VHA clinic, and help other veteran patients get care.

> *Your clinician can actually be seeing somebody that's really in need instead of doing a basic maintenance check.*

## Clinician Surveys

Of 22 clinicians interviewed, 20 (87%) participated in the survey, 14 (64%) of which were fully complete. Of the 20 respondents, 6 were MD/DOs, 7 were advanced practice providers (APPs), 6 were registered nurses (RNs), and 1 was a medical instrument technician (Table 3). Ten self-identified as female and 6 self-identified as non-White. Almost half of the respondents had been working at their current VHA cardiology clinic for >10 years. All clinicians were focused on CIED-related care and were not serving as patients' primary cardiology clinician.

The most commonly reported scheduling frequency for routine in-person ICD and pacemaker evaluations was every 12 (range 4 - 12) months, used by 72% (n=13) and 83% (n=15) of clinicians, respectively (Table 4).

Seven (39%) clinicians reported using an RM-first strategy for some patients. Sixteen (89%) thought this strategy would improve veteran convenience by reducing appointments and travel time. Six (33%) expected it would enable more care for other patients with heart rhythm disorders.

However, 12 (63%) clinicians were concerned about a reduction in the quality of veteran care and 10 (53%) about veteran-perceived abandonment. Fifteen (83%) respondents were confident that an RM-first strategy was as effective as RM with in-office visits regarding cardiovascular outcomes, while 3 (17%) were not. Seven (39%) expected an RM-first strategy would benefit their clinic, 7 (39%) were undecided, and 4 (22%) thought it would not.

**Table .** Clinician characteristics and perspectives on remote monitoring (RM)–first strategy.

| Characteristic | | Values, n (%) |
|---|---|---|
| Title (n=20) | | |
| | Advanced practice provider | 7 (35) |
| | Medical instrument technician | 1 (5) |
| | Registered nurse | 6 (30) |
| | Physician | 6 (30) |
| Time worked with current VHA[a] cardiology clinic (n=20) | | |
| | <1 year | 0 (0) |
| | 1 - 5 years | 8 (40) |
| | 6 - 10 years | 3 (15) |
| | >10 years | 9 (45) |
| Adjustment to CIED[b] care schedule if the patient does not want routine in-person CIED checks or misses an in-person check (n=19)[c] | | |
| | Adjust the RM transmission schedule | 3 (16) |
| | Reduce the frequency of in-person device checks | 5 (26) |
| | Offer video visit paired with RM as an alternative | 2 (11) |
| | Offer a telephone visit paired with RM as an alternative | 9 (43) |
| | Other: encourage rescheduling an in-person visit | 3 (16) |
| Current use of RM-first strategy for any patients (n=18) | | |
| | Yes | 7 (39) |
| | No | 11 (61) |
| Benefits for RM-first strategy (n=18)[c] | | |
| | Veteran convenience in reducing appointments and travel time | 16 (89) |
| | Better use of clinic space | 7 (39) |
| | Ability to see other patients with heart rhythm disorders | 6 (33) |
| Concerns about an RM-first strategy (n=18)[c] | | |
| | Changes to payment structure or relative value units | 2 (11) |
| | Reduction in quality of veteran care | 12 (63) |
| | Veteran patient impression of abandonment | 10 (53) |
| | Reducing veteran contact with the VHA | 9 (47) |
| Confidence that an RM-first strategy is as effective as RM + in-office evaluations for cardiovascular outcomes (n=18) | | |
| | Not at all confident | 3 (17) |
| | Somewhat confident | 10 (56) |
| | Confident | 3 (17) |
| | Somewhat more confident | 1 (5) |
| | Very confident | 1 (5) |
| Would an RM-first strategy help your clinic? (n=18) | | |
| | Yes | 7 (39) |
| | No | 4 (22) |
| | Undecided | 7 (39) |

[a]VHA: Veterans Health Administration.

[b]CIED: cardiovascular implantable electronic device.

[c]Participants able to select multiple responses.

**Table .** Current frequency of routine in-person evaluations and remote transmission reviews reported by clinicians.

| | For patients with implantable cardioverter-defibrillators, n (%) | For patients with pacemakers, n (%) |
|---|---|---|
| Frequency of routine in-person evaluation (n=18 clinicians) | | |
| 4 months | 1 (6) | 0 (0) |
| 6 months | 4 (22) | 2 (11) |
| 10 months | 0 (0) | 1 (6) |
| 12 months | 13 (72) | 15 (83) |
| Frequency of transmission review without an in person visit (n=14 clinicians) | | |
| 3 months | 4 (29) | 4 (29) |
| 5 months | 1 (7) | 0 (0) |
| 10 months | 0 (0) | 1 (7) |
| 12 months | 1 (7) | 1 (7) |
| Not applicable | 8 (57) | 8 (57) |

## Clinician Interviews

### *Overview*

Most interviewed clinicians were open to RM-first care, although some were not, and a few had no preference. Although many were hesitant, they still expected that RM-first care represented the future.

Many clinicians already had experience with RM-first care during the COVID-19 pandemic and noted that it reduced veteran travel time and clinician visit burden, but patient RM connectivity was a challenge. Most clinicians and facilities had returned to the prepandemic model of CIED care. Barriers and facilitators to RM-first care described by clinicians are in Table 2.

### *Barriers to RM-First Care*

#### Lacking Routine In-Person Assessment

The most cited barrier by clinicians was that the benefits of routine in-person assessment during CIED clinic visits would not be available. These concerns ranged from a general sense that an in-person assessment was safer for patients, particularly for patients with greater complexity, such as those with advanced heart failure, to specifically valuing the physical examination and opportunity for in-person medication reconciliation. As a medical instrument technician stated,

> *If we cannot assess their condition in-person, then we may find flags later that are really big issues and then we have to adjust everything.*

These concerns could also be related to missing important CIED information, including the occasional need for reprogramming.

#### Reducing Veteran Contact With VHA

Another clinician-cited barrier was that an RM-first approach would lead to a reduction in veteran contact with the VHA, which could potentially leave patients perceiving abandonment. As one RN stated,

> *In-person visits are the expectation for many patients, so they could feel abandoned.*

A physician discussed the importance of the rapport built during routine in-person CIED visits,

> *Face-to-face interactions with patients and doctors [are] important for rapport. Just putting your hand on them can make your relationship and their comfort with you better.*

Some clinicians expressed concern that patients would be lost to follow-up without in-person visits because device clinic visits are used to ensure that patients have other routine cardiology follow-up scheduled. As a physician stated,

> *Patients always get lost to follow-up so it's nice to have one more place to get eyes on them.*

#### Clinic Operations–Related Changes

Clinicians anticipated the need for operational changes to their clinic, including ensuring a reliable tracking system for patients not being seen in person to prevent patients from being lost to follow-up. As an APP stated,

> *I don't know that we have a system in place for the clinic as a whole to track things… between the device nurse, the provider and the EP nurse navigator [we would need] to develop some sort of tracking system.*

Clinicians also perceived a need for time to review more remote transmissions if patients were not receiving routine in-person device clinic evaluations. As an APP shared,

> *Definitely more time on the nursing side to… get them [remote transmissions] processed into the charting system.*

Some felt that without an in-person visit, at least an annual review of the patient's data would be important.

> *I would still want a yearly review… I would go through it with a fine-toothed comb.*

Finally, there were concerns surrounding the loss of device clinician skills if patients were no longer routinely attending in-person visits, particularly for training new staff. As one RN shared,

> *As self-taught on remote monitoring, we will get rusty on our skills… The learning curve is pretty steep… to feel comfortable to perform an interrogation independently. In-person clinic follow up is our only way of training… If we went remote-only, we would have no way of both training new staff and keeping current comfortable. Then when we would need to see patients, we would be at a severe disadvantage.*

#### Technological Difficulties

Interviewees noted that an RM-first approach placed increased importance on RM technology and some worried that veterans and clinicians may experience technological difficulties, particularly because RM adherence and connectivity were essential. As an RN stated,

> *The tech is the stumbling block because it's hard to troubleshoot the home monitor when it's not working. Then you have to make them come in and some would not want to come after not coming for a while.*

### Benefits of RM-First Care

#### *Reduced Veteran Travel Burden*

Interviewees emphasized reduced veteran travel burden—including reduced travel time, cost, and weather-related issues. As an RN stated,

> *[RM-first care] would be good for those patients who travel 200+ miles for 15-minute visits.*

Similarly, an electrophysiologist stated,

> *Some drive more than 100 miles to get here... Winter storms are another example when it is dangerous to travel.*

An RN explained that some patients have difficulty arranging transportation and are unable to drive themselves to clinic visits,

> *Some patients have 4 hours travel to our clinic…Staying home and only coming in for reprogramming needs would be useful. Cost has gone up as well, with fuel prices, being on the road and eating out. There are not great DAV transportation options. A lot of problems finding van drivers.*

Finally, a few clinicians thought that RM-first care may make some patients more likely to engage in CIED care. As one electrophysiologist noted,

> *Some patients really turn off about having to come in. There are some who are more likely to engage through remote monitoring only.*

#### *Optimization of Clinic Staff Time and Clinic Space*

Another potential benefit of RM-first care was that it could optimize clinic staff time and often-limited outpatient clinic space. As 1 physician described,

> *It would offload clinics, that's [in-person CIED visits] a lot of work that APPs do. They could devote more time to a multitude of other tasks.*

The time could be used to evaluate other patients with heart rhythm disorders waiting for care, explained an APP,

> *Downsizing device clinic space could increase in-person arrhythmia clinic space.*

#### *Increased Time to Review Routine Remote Transmissions and Improve RM Adherence*

Interviewees also mentioned that an RM-first care model could increase staff time to review routine remote transmissions and support RM adherence. One APP explained,

> *Some of those remote transmissions are over 100 pages long. There are days when I get 10 or more device alerts and it takes time to go through EGMs (intracardiac electrograms) and not missing anything. It would provide more time on the nursing side.*

#### *No Concern About Relative Value Unit Workload Credit*

Finally, most clinicians thought there would be no issue with relative value units (RVUs) when transitioning to an RM-first model. As an RN said,

> *No [concerns regarding RVUs]. ... Sometimes you get more RVUs reviewing patients' remote transmissions. You can do a note for addressing a missed transmission. People need to know the benefit of reviewing more remote transmissions.*

### Implementation of RM-First Care

Clinicians thought that patients who were the best candidates for RM-first care were those without cardiac resynchronization therapy (CRT) devices who were adherent to RM, clinically stable and noncomplex, not device-dependent, not having frequent arrhythmias, good communicators, and facile with technology. One APP explained,

> *There is a certain population that would be appropriate. Younger, less comorbidities, low pacing burdens, that sort of thing. Knowledgeable and familiar with RM.*

Many clinicians expected the decision about appropriateness for an RM-first strategy would initially be determined by the patient's clinician, as an APP explained,

> *Anyone that the provider deems appropriate. It will be joint decision-making between the patient and the provider. We will talk with them and assess what their goals are, and as long as they understand that based on remote monitoring they would still have to come into the clinic if clinically indicated.*

When asked how an RM-first care model should be evaluated for success, most clinicians thought patient satisfaction should

be a key indicator, along with patient RM adherence. As an APP said,

> *Adherence to remote monitoring. I think you would want adherence over 95%. How are the Vets feeling about it, are they satisfied? Surveys. A lot of Vets would be amenable.*

Respondents also thought it would be important to ensure there was no increase in adverse outcomes or rhythm issues not being identified promptly.

> *Prove that there are no greater adverse cardiac outcomes. I will always be more conservative with my Veteran patients and wary of big changes in care.*

Respondents also discussed potential time savings with an RM-first approach. As an RN said,

> *Measure time savings of remote monitoring.*

Many interviewees also noted that monitoring for missed RM transmissions would be central for a new RM-first care model, but most already had a process in place for doing so. One APP explained,

> *We would follow the same scheduling tracking system we have now. It's basically a log by manufacturer and when they were last seen.*

## Discussion

### Principal Results

The 2023 HRS expert consensus statement introduced "alert-based remote monitoring," defined as "a combination of continuous connectivity with clinic visits that are prompted only by the detection of actionable events," [2] which provides the basis for the RM-first care model that we discussed with veterans and clinicians. Both expressed interest in this model of CIED care and cited the benefit of reducing patient travel burden and enabling clinical bandwidth to care for other patients. However, patients sometimes preferred in-person evaluations (generally for non-CIED related medical reasons and the patient-clinician relationship), and some expressed concerns regarding technological issues with RM. Given the VHA's central RM infrastructure that reviews all remote transmissions, VHA is well-positioned to implement and study this care model, which could inform other health systems and clinicians about the context of implementing RM-first care. Indeed, most clinicians expected that RM-first would ultimately become the standard of care for CIED management.

### Comparison With Prior Work

There is often substantial lag in implementing research and consensus recommendations into clinical practice, including inertia in initiating new care models [21,22]. Reasons for such inertia include overestimation of existing care as well as lack of practice organization to achieve therapeutic goals [22]. Providing patient and clinician education and support when implementing an RM-first care model will be important to overcome inertia, leverage facilitators, and surmount barriers.

### Strategies to Overcome Barriers in Implementation

Some patients worried about the quality of RM. To address this, patient-centered RM education should be provided before transitioning to RM-first care and emphasize to patients that any actionable findings on RM will prompt appropriate clinical actions, sometimes including in-person evaluations. Additionally, for patients to qualify for this care strategy, they need to be consistently and continuously connected to RM so clinically actionable events can be identified promptly. Thus, patients should be educated about ensuring RM connectivity and troubleshooting strategies based on their specific transmitter. Patients and clinicians also raised concerns regarding the loss of the in-person relationship and the inability to perform in-person assessment, such as a physical examination. To address this, device clinicians should ensure that patients have regular follow-ups with their general cardiologist or electrophysiologist (as appropriate) or at least routine primary care, and that the device clinic is not their primary source of cardiology care.

Clinicians also noted a potential increased risk of patients being lost to follow-up. Clinics must have a method of tracking patients outside of in-person visits and ensuring RM adherence [16]. Patients who become disconnected from RM will require in-person evaluation. Finally, patients and clinicians raised concerns about technical comfort with troubleshooting home monitors and RM adherence, which requires a high workload [23]. To alleviate this burden, postcard reminders that recommend patients contact their CIED manufacturer for assistance have been shown to increase RM adherence, without burdening clinicians [24]. Additionally, sending informational text messages to recently disconnected patients can improve RM adherence [25].

### Benefits of Implementation

Although there are several barriers to be addressed, the RM-first care model has the potential to provide many improvements for patients and clinicians. With the growing potential of digital health technology in cardiovascular medicine [26], the lessons from our study have broad applicability but it will be critical to ensure that an RM-first care model, as with any virtual care modality, is implemented equitably [27,28]. Reduced patient travel burden is particularly important for patients who live in rural locations. From a reimbursement perspective, while VHA is a single-payer, other health care payers would need to adopt novel reimbursement strategies for RM that facilitate sustainable and cost-effective CIED follow-up care [2,29,30]. Finally, a reduction in unnecessary device-related clinic visits will allow clinicians to see other patients with heart rhythm disorders and reduce wait times, which may result in higher-value care, particularly given the shortage of cardiovascular health professionals [31]. An RM-only model has been successfully implemented at a large clinic in Italy since the COVID-19 pandemic and was associated with time savings for clinicians and patients with no increase in adverse clinical outcomes [32]. Further, although not currently available, if remote reprogramming is demonstrated to be safe and feasible to implement, it could further reduce the need for in-person visits

and could improve patient perceptions around an RM-first care model.

### Limitations

Our study should be considered in the context of its limitations. First, although we studied a single health system with specific patient population demographics (more often rural, predominantly White, and predominantly male) and clinicians providing care in an integrated health care delivery system, the Veterans Affairs National Cardiac Device Surveillance Program (VANCDSP) centrally monitors more than 64,000 veterans with CIEDs, making VHA well-positioned to implement and evaluate RM-first care. Future studies should evaluate other patient populations, which would help to assess the transferability of our findings. Second, although this was a national study, our results represent a limited number of both patient and clinician perspectives. However, qualitative methods intentionally provide granular data from smaller numbers of participants, patients were randomly selected, and our methodology provided detailed information on perspectives from clinicians across the United States. Third, interviews were conducted while new HRS consensus was released in draft form [2], so questions were modified partway through the interview process, and the ideas being introduced were new; patients and clinicians may feel differently when they have had more time to assimilate the recommendations. We did not inform patients about the additional safety offered by consistent and continuous RM connectivity. Fourth, we did not interview patients who were new or nonadherent to RM. Fifth, we did not have participant validation of our findings. Sixth, this study's team represented an institution (VANCDSP) with some influence on both patient care and clinical support. While it was not apparent in the review of interview recordings or transcripts, this power dynamic may have incentivized veteran patients and clinicians to speak more favorably of the VANCDSP or caused interviewees to present their care or their patient's existing care in a more favorable light. Finally, this study represents patient and clinician expectations of RM-first care, instead of their views based on experience; as RM-first is implemented in the future, patient and clinician perceptions on barriers and facilitators to this care model should be evaluated.

### Conclusions

Both patients and CIED clinicians experienced in RM within the VHA were open to an RM-first care model that reduces in-person visits but conveyed barriers about solely relying on RM and possible diminution of the patient-clinician relationship. Implementation of new RM recommendations will require attention to these perceptions and prioritization of patient-centered approaches.

## Acknowledgments

We thank Gary Tarasovsky, BS, and Gregory Rohrbach, DNP, for helping us to identify patients and clinicians for interviews. This work was funded by the Department of Veterans Affairs (VA), Veterans Health Administration (VHA), Office of Rural Health [ORH], NOMAD #PROJ-03669, VA Health Systems Research (1IK2HX003357), and VA Health Systems Research Quality Enhancement Research Initiative (I50 HX003266). The contents do not represent the views of the Department of Veterans Affairs or the United States government.



## Data Availability

Data outside of those reported in this paper are not applicable to data sharing due to privacy constraints.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Final clinician and veteran interview guides and clinician survey.
[DOCX File, 41 KB - cardio_v9i1e66215_app1.docx ]

## References

1. Slotwiner D, Varma N, Akar JG, et al. HRS Expert Consensus Statement on remote interrogation and monitoring for cardiovascular implantable electronic devices. Heart Rhythm 2015 Jul;12(7):e69-e100. [doi: 10.1016/j.hrthm.2015.05.008] [Medline: 25981148]
2. Ferrick AM, Raj SR, Deneke T, et al. 2023 HRS/EHRA/APHRS/LAHRS expert consensus statement on practical management of the remote device clinic. Heart Rhythm 2023 Sep;20(9):e92-e144. [doi: 10.1016/j.hrthm.2023.03.1525] [Medline: 37211145]
3. Akar JG, Bao H, Jones PW, et al. Use of remote monitoring is associated with lower risk of adverse outcomes among patients with implanted cardiac defibrillators. Circ Arrhythm Electrophysiol 2015 Oct;8(5):1173-1180. [doi: 10.1161/CIRCEP.114.003030] [Medline: 26092577]

4. Varma N, Piccini JP, Snell J, Fischer A, Dalal N, Mittal S. The relationship between level of adherence to automatic wireless remote monitoring and survival in pacemaker and defibrillator patients. J Am Coll Cardiol 2015 Jun 23;65(24):2601-2610. [doi: 10.1016/j.jacc.2015.04.033] [Medline: 25983008]
5. Mittal S, Piccini JP, Snell J, Prillinger JB, Dalal N, Varma N. Improved survival in patients enrolled promptly into remote monitoring following cardiac implantable electronic device implantation. J Interv Card Electrophysiol 2016 Aug;46(2):129-136. [doi: 10.1007/s10840-016-0112-y] [Medline: 26860839]
6. Piccini JP, Mittal S, Snell J, Prillinger JB, Dalal N, Varma N. Impact of remote monitoring on clinical events and associated health care utilization: a nationwide assessment. Heart Rhythm 2016 Dec;13(12):2279-2286. [doi: 10.1016/j.hrthm.2016.08.024] [Medline: 27544748]
7. Guédon-Moreau L, Lacroix D, Sadoul N, et al. A randomized study of remote follow-up of implantable cardioverter defibrillators: safety and efficacy report of the ECOST trial. Eur Heart J 2013 Feb;34(8):605-614. [doi: 10.1093/eurheartj/ehs425] [Medline: 23242192]
8. Timmermans I, Meine M, Szendey I, et al. Remote monitoring of implantable cardioverter defibrillators: patient experiences and preferences for follow-up. Pacing Clin Electrophysiol 2019 Feb;42(2):120-129. [doi: 10.1111/pace.13574] [Medline: 30536931]
9. Watanabe E, Yamazaki F, Goto T, et al. Remote management of pacemaker patients with biennial in-clinic evaluation: continuous home monitoring in the japanese at-home study: a randomized clinical trial. Circ Arrhythm Electrophysiol 2020 May;13(5):32342703. [doi: 10.1161/CIRCEP.119.007734] [Medline: 32342703]
10. García-Fernández FJ, Osca Asensi J, Romero R, et al. Safety and efficiency of a common and simplified protocol for pacemaker and defibrillator surveillance based on remote monitoring only: a long-term randomized trial (RM-ALONE). Eur Heart J 2019 Jun 14;40(23):1837-1846. [doi: 10.1093/eurheartj/ehz067] [Medline: 30793735]
11. Varma N, Love CJ, Michalski J, Epstein AE, TRUST Investigators. Alert-based ICD follow-up: a model of digitally driven remote patient monitoring. JACC Clin Electrophysiol 2021 Aug;7(8):976-987. [doi: 10.1016/j.jacep.2021.01.008] [Medline: 33640345]
12. Muniyappa AN, Raitt MH, Judson GL, et al. Factors associated with remote monitoring adherence for cardiovascular implantable electronic devices. Heart Rhythm 2022 Sep;19(9):1499-1507. [doi: 10.1016/j.hrthm.2022.04.025] [Medline: 35500792]
13. Office of Rural Health. Rural Veterans. US Department of Veterans Affairs. URL: https://www.ruralhealth.va.gov/aboutus/ruralvets.asp [accessed 2025-03-19]
14. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
15. Dhruva SS, Raitt MH, Munson S, et al. Barriers and facilitators associated with remote monitoring adherence among veterans with pacemakers and implantable cardioverter-defibrillators: qualitative cross-sectional study. JMIR Cardio 2023 Nov 21;7:e50973. [doi: 10.2196/50973] [Medline: 37988153]
16. Rotering TL, Hysong SJ, Williams KE, Raitt MH, Whooley MA, Dhruva SS. Strategies to enhance remote monitoring adherence among patients with cardiovascular implantable electronic devices. Heart Rhythm O2 2023 Dec;4(12):794-804. [doi: 10.1016/j.hroo.2023.11.002] [Medline: 38204458]
17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
19. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
20. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292(114523):114523. [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
21. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med 2011 Dec;104(12):510-520. [doi: 10.1258/jrsm.2011.110180] [Medline: 22179294]
22. Phillips LS, Branch WT, Cook CB, et al. Clinical inertia. Ann Intern Med 2001 Nov 6;135(9):825-834. [doi: 10.7326/0003-4819-135-9-200111060-00012] [Medline: 11694107]
23. Harvey M, Seiler A. Challenges in managing a remote monitoring device clinic. Heart Rhythm O2 2022 Feb;3(1):3-7. [doi: 10.1016/j.hroo.2021.12.002] [Medline: 35243430]
24. McLaughlin MM, Raitt MH, Tarasovsky G, Whooley MA, Dhruva SS. Informational postcards increase engagement with remote monitoring among veterans with pacemakers and implantable cardioverter-defibrillators: a stepped-wedge randomized controlled trial. J Gen Intern Med 2024 Feb;39(Suppl 1):87-96. [doi: 10.1007/s11606-023-08478-9] [Medline: 38252247]
25. Durand J, Bonnet JL, Lazarus A, Taieb J, Rosier A, Mittal S. Using technology to improve reconnection to remote monitoring in cardiac implantable electronic device patients. Cardiovasc Digit Health J 2024 Feb;5(1):1-7. [doi: 10.1016/j.cvdhj.2023.11.020] [Medline: 38390582]

26. Varma N, Han JK, Passman R, et al. Promises and perils of consumer mobile technologies in cardiovascular care: JACC scientific statement. J Am Coll Cardiol 2024 Feb 6;83(5):611-631. [doi: 10.1016/j.jacc.2023.11.024] [Medline: 38296406]
27. Tisdale R, Der-Martirosian C, Yoo C, Chu K, Zulman D, Leung L. Disparities in video-based primary care use among veterans with cardiovascular disease. J Gen Intern Med 2024 Feb;39(Suppl 1):60-67. [doi: 10.1007/s11606-023-08475-y] [Medline: 38252244]
28. Zulman DM, Wong EP, Slightam C, et al. Making connections: nationwide implementation of video telehealth tablets to address access barriers in veterans. JAMIA Open 2019 Oct;2(3):323-329. [doi: 10.1093/jamiaopen/ooz024] [Medline: 32766533]
29. Yang S, Stabenau HF, Kiernan K, Diamond JE, Kramer DB. Clinical utility of remote monitoring for patients with cardiac implantable electrical devices. J Interv Card Electrophysiol 2023 Jun;66(4):961-969. [doi: 10.1007/s10840-022-01406-7] [Medline: 36327060]
30. Mecklai K, Smith N, Stern AD, Kramer DB. Remote patient monitoring - overdue or overused? N Engl J Med 2021 Apr 15;384(15):1384-1386. [doi: 10.1056/NEJMp2033275] [Medline: 33853209]
31. Fry ETA. Resigned to the “Great Resignation?”. J Am Coll Cardiol 2022 Jun 21;79(24):2463-2466. [doi: 10.1016/j.jacc.2022.05.004] [Medline: 35710197]
32. Miracapillo G, Addonisio L, De Sensi F, et al. Switching to a 100% remote follow-up of implantable cardiac electronic devices: organizational model and results of a single center experience. J Cardiovasc Electrophysiol 2024 Aug;35(8):1548-1558. [doi: 10.1111/jce.16328] [Medline: 38818537]


## Abbreviations

**APP:** advanced practice provider
**CIED:** cardiovascular implantable electronic device
**HRS:** Heart Rhythm Society
**ICD:** implantable cardioverter-defibrillator
**RM:** remote monitoring
**RN:** registered nurse
**RVU:** relative value unit
**VA:** Veterans Affairs
**VANCDSP:** Veterans Affairs National Cardiac Device Surveillance Program
**VHA:** Veterans Health Administration


*Edited by A Coristine; submitted 06.09.24; peer-reviewed by J Edwards, YM Hwang; revised version received 13.02.25; accepted 13.02.25; published 04.04.25.*

*Please cite as:*
*Kratka A, Rotering TL, Munson S, Raitt MH, Whooley MA, S Dhruva S*
*Patient and Clinician Perspectives on Alert-Based Remote Monitoring–First Care for Cardiovascular Implantable Electronic Devices: Semistructured Interview Study Within the Veterans Health Administration*
*JMIR Cardio 2025;9:e66215*
*URL: https://cardio.jmir.org/2025/1/e66215*
*doi:10.2196/66215*

# Novel Virtual Reality Intervention for Stress Reduction Among Patients With or at Risk for Cardiovascular Disease: Mixed Methods Pilot Study

Katherine E Makaroff[1], MSHS, MD; Christopher Van[1]; Vincent Grospe[2], BSci; Lynae Edmunds[1], BA; Marcella A Calfon-Press[1], MD, PhD; Karol E Watson[1], MD, PhD; Tamara Horwich[1], MS, MD

[1]David Geffen School of Medicine at University of California, Los Angeles, 10833 Le Conte Avenue, Los Angeles, CA, United States

[2]University of California, San Francisco School of Medicine, San Francisco, CA, United States

**Corresponding Author:**
Tamara Horwich, MS, MD
David Geffen School of Medicine at University of California, Los Angeles, 10833 Le Conte Avenue, Los Angeles, CA, United States

## Abstract

**Background:** Virtual reality (VR) has emerged as a promising, low-risk strategy to manage many forms of psychological stress and may be a modality to improve cardiovascular health. Recent scientific statements on the mind-heart-body connection call for better adherence to psychological screening and adoption of more holistic “behavioral cardiology” interventions that improve the overall health of patients with or at risk for cardiovascular disease (CVD).

**Objective:** The aim of this study is to assess safety and preliminarily explore how a VR experience can aid in stress reduction among patients with or at risk for CVD.

**Methods:** A convergent mixed methods approach was used for this single-arm prospective pilot study. In total, 20 patients were recruited from the University of California Los Angeles adult cardiology clinics and cardiac rehabilitation. Surveys and physiologic parameters were collected before, during, and after a 30-minute VR experience aimed at relaxation. The primary outcome was the State-Trait Anxiety Inventory-State (STAI-S) scale. They participated in a 90-minute visit, during which they completed surveys, including the STAI-S scale, before and after a 30-minute VR experience. Physiological parameters were also collected before, during, and after the experience. Visits concluded with semistructured interviews analyzed with inductive thematic analysis to add depth and nuance to our analysis.

**Results:** STAI-S scale scores after the VR experience were significantly decreased from baseline (median 31, IQR 28-38 vs median 24, IQR-29.25; $P<.001$). Verbal feedback revealed that participants experienced a relaxing sense of “distance from stress” moderated by unexpected, intense audiovisual components. Heart rate significantly decreased (mean 73, SD 8 vs mean 67, SD 6 beats per minute; $P<.001$), while blood pressure (mean systolic 128, SD 14 vs mean systolic 129, SD 18 mm Hg; $P=.75$ and mean diastolic 79, SD 9 vs mean diastolic 80, SD 10 mm Hg; $P=.60$) and galvanic skin response (mean 0.74, SD 0.89 vs mean 0.70, SD 0.57 microsiemens; $P=.45$) remained the same. Changes in heart rate variability parameters were consistent with increased vagal tone over time but were only statistically significant at certain time points. Survey results and interviews generally indicated safety, tolerability, and openness to using VR again.

**Conclusions:** This sample of patients with CVD or risk of CVD had above-average stress, consistent with epidemiological data; the statistically and clinically significant decrease in subjective perception of stress partially converged with physiologic data. Overall, the VR intervention was a safe and feasible stress reduction method. Future research is needed to evaluate the effectiveness of this immersive therapy in reducing cardiovascular risk profiles.

***(JMIR Cardio 2025;9:e66557)*** doi:10.2196/66557

**KEYWORDS**

heart disease risk factors; stress reduction; digital health; pilot; virtual reality; stress; risk; cardiovascular disease; CVD; behavioral cardiology; mixed methods; cardiology; cardiac rehabilitation; survey; blood pressure; heart rate

## Introduction

### Background

Western medicine has traditionally treated the heart and mind as separate entities. However, emerging data point to a powerful "mind-heart-body connection" in which all are interconnected and interdependent [1-3]. Research has clearly demonstrated that positive psychological states are associated with a lower risk of cardiovascular disease (CVD) and mortality, while negative psychological factors such as chronic stress, anxiety, and depression can negatively impact cardiovascular health [1-10]. Further, chronic stress is associated with a 40% - 50% increase in the risk of coronary artery disease [4]. Studies demonstrate that persistent psychological distress, including anxiety and depression, is an independent predictor of morbidity and mortality in those with established CVD [11-13]. Data from the international case-control, also known as the INTERHEART study (effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries), show that psychosocial stressors account for 33% of the risk for myocardial infarction [10,14]. Psychological health in patients with or at risk for CVD represents an opportunity for risk reduction. The American Heart Association's 2021 Scientific Statement on psychological health, well-being, and the mind-heart-body connection called for better adherence to psychological screening measures and adoption of more holistic "behavioral cardiology" interventions that improve overall health in patients with or at risk for CVD [2].

To aid in several health care challenges, many physicians and patients are harnessing the power of emerging technologies. Virtual reality (VR) is a technology that provides an immersive experience of a computer-generated, 3D image or environment using a head-mounted display. VR has presented a low-risk way to interrupt the brain's "default mode network" (DMN), which is a particular set of brain structures that underlie the negative mental states marked by worry, rumination, and stress [15-18]. The DMN is also responsible for the "baseline buzzing" and drive for the mind to wander or "forage" for new information even when trying to relax or think about nothing at all [18]. Further, it has been shown that people become unhappier the longer time their minds spend wandering [19]. VR has now demonstrated promise as a treatment modality for anxiety, phobias, depression, autism, and posttraumatic stress disorder as well as a way to aid in meditation [20-29]. Thus, VR is a potentially powerful tool to target stress reduction in patients with CVD.

### Aims and Research Question

The aim of this study was to determine whether stress levels could be reduced immediately after experiencing a novel VR intervention in patients sampled from the University of California Los Angeles (UCLA) cardiology clinics and cardiac rehabilitation (CR) program. This project directly addressed the "heart-mind-body" connection by using an innovative VR intervention aimed at reducing stress in those with CVD. This aim was addressed through a convergent mixed methods design: this question was examined quantitatively by changes in (1) subjective patient-reported stress levels on a validated survey, (2) blood pressure (BP), (3) galvanic skin response (GSR), (4) heart rate (HR), and (5) heart rate variability (HRV). Semistructured interviews performed after the VR experience were analyzed using inductive thematic analysis to understand participant experiences and the potential therapeutic value of the VR experience. The quantitative and qualitative data were examined in parallel and integrated to look for areas of convergence or divergence across the data.

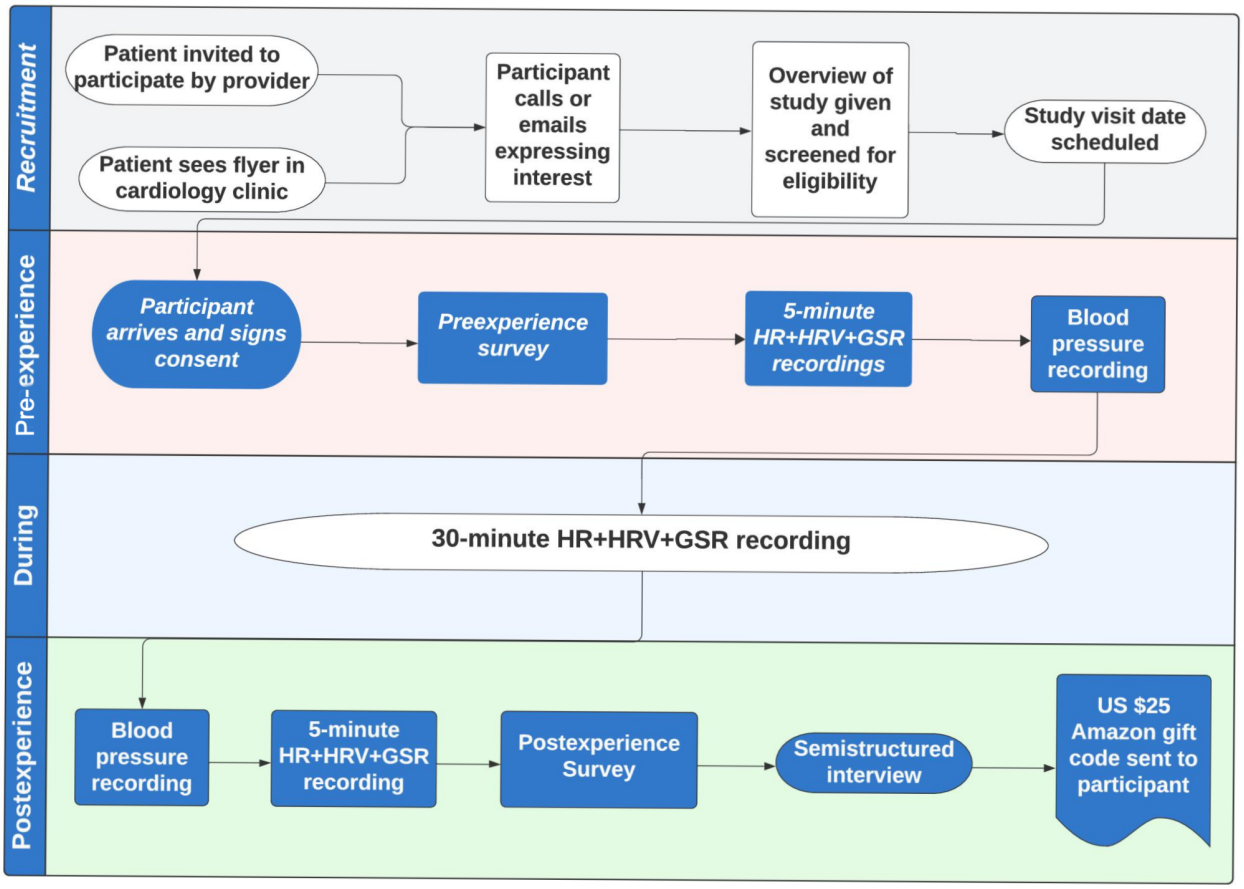
## Methods

### Study Design

According to a framework proposed by a multidisciplinary group of experts for assessing immersive therapeutics [30], this study blends features of a VR1 study, which focuses on developing content in partnership with patients, and VR2 study, which tests treatment feasibility, acceptability, and tolerability. Going through this piloting phase is vital prior to designing subsequent studies comparing clinical outcomes between one group receiving the VR treatment and another receiving a control treatment (ie, VR3). A convergent mixed methods approach was used for this single-arm prospective pilot study. Figure 1 displays a study procedure flowchart.

Eligible participants were recruited to attend 1 in-person study visit lasting ~90 minutes. Participants were asked to sit and complete informed consent before proceeding to complete a pre-experience survey that took about 5 to 10 minutes to complete. It included the Perceived Stress Scale-10 (PSS-10) [31-34], 20-item State-Trait Anxiety Inventory-State scale (STAI-S) [35-40], and level of experience with various stress reduction methods. It also asked about exercise, caffeine, and sleep both generally and on the day of the study. Next, participants were connected to equipment, and a 5-minute HR and HRV recording was collected using the well-validated Polar H10 HR monitor [41,42], a commercial chest-based electrocardiogram strap considered among researchers as the standard for accurately quantifying cardiovascular metrics after multilead electrocardiograms [43-45]. The HR monitor was connected via Bluetooth to the validated third-party smartphone app EliteHRV [46] (while the patient rested quietly in a semirecumbent position). A 5-minute GSR [47-49] recording was taken simultaneously using the validated NeuLog sensor and software [50]. BP was recorded with an Omron 3 series digital BP monitor and cuff immediately before and after the VR experience (delivered through a Meta Quest 2 VR headset). During the experience, participants remained in the semirecumbent position, while HR, HRV, and GSR were measured continuously. A final 5-minute postexperience HR, HRV, and GSR recording was taken. Participants then completed the postexperience survey consisting of the STAI-S scale, Immersive Tendencies Questionnaire (ITQ) [51], Simulation Sickness Questionnaire (SSQ) [52,53], overall rating of VR intervention, level of prior experience with VR, and demographics. The study visit concluded with a 10 - to 15-minute semistructured interview. Participants were emailed a US $25 Amazon e-gift code.

**Figure 1.** Study procedure flowchart. GSR: galvanic skin response; HR: heart rate; HRV: heart rate variability.




## Ethical Considerations

The study was approved by the UCLA Institutional Review Board (#21‐000705), and the ClinicalTrials.gov registry number was NCT0498465. Written informed consent was obtained. The privacy and confidentiality of research participants' data and identity were maintained. Participants were compensated with US $25 Amazon gift cards.

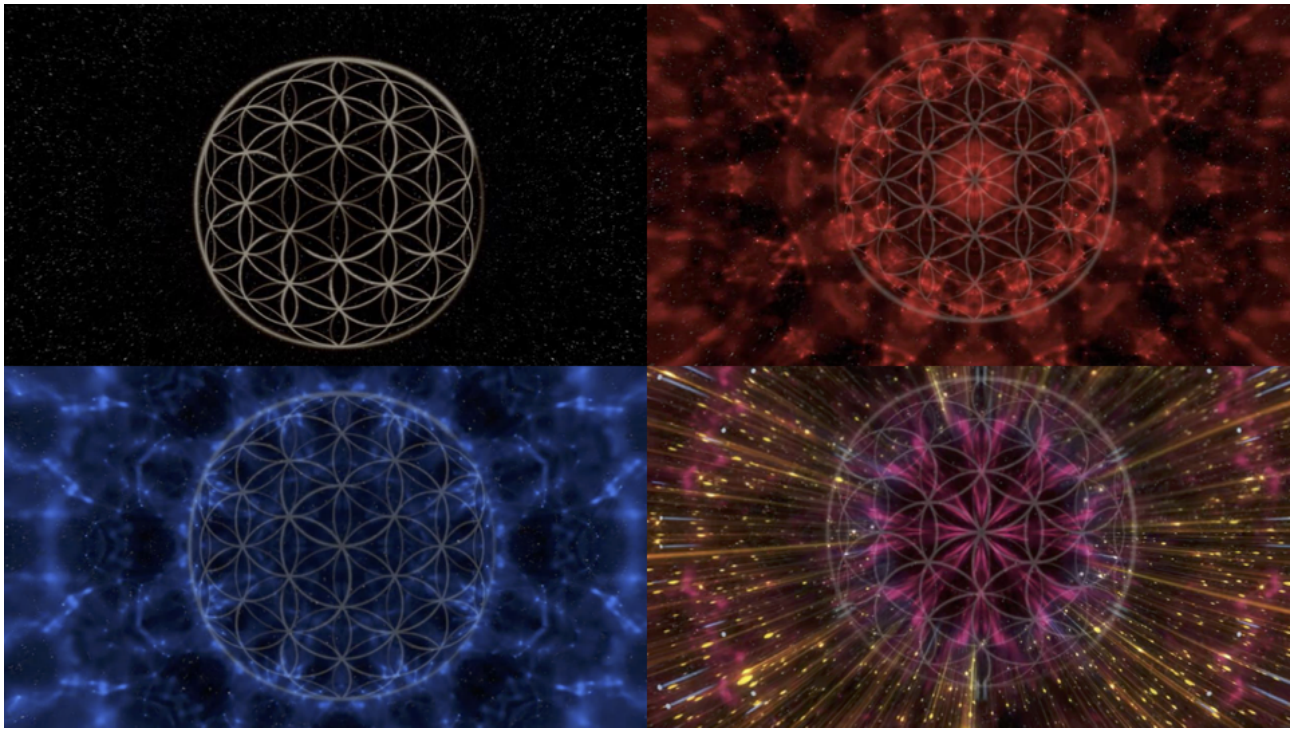
## Recruitment

We recruited 20 patients from the UCLA cardiology outpatient clinics and CR program who were 18 years and older of age, English speaking (rationale for English only: surveys and consent are not translated into languages other than English and study team members are primarily English speakers, there was no funding to provide translation or medical interpreters during participation), and able to give informed consent. Patients were excluded due to the presence of conditions that interfere with VR use, including history of seizure, facial injury, significant hearing or visual impairment, individuals with dangerous unstable arrhythmias (ventricular tachycardia or fibrillation) or a myocardial infarction in the past 4 weeks, individuals in acute decompensated heart failure, those with factors known to contribute to cybersickness, including postural instability or motion sickness. Potentially eligible participants self-referred or were referred to a research coordinator by the primary investigator after an initial chart review. These patients were screened over the phone to confirm eligibility and assess interest.

## VR Experience

This 30-minute experience developed by Harmony Media Company is an immersive therapy created to lower stress and anxiety. The experience was not designed specifically for those with CVD, but rather for use in a more general population. User feedback regarding the experience was obtained informally from healthy adults at several points during development. It delivers a proprietary combination of colorful and fractal sacred geometric visual effects synchronized with a nonverbal, binaural audioscape that aims to exert its therapeutic effect by disrupting the DMN of brain signaling (Figure 2). The content is comprised of predefined (noninteractive) primary and background looping visuals, animated textures, and specific prerendered sequences allowing for predictability, measurability, and performance stability. The focus area (approximately 80% of the visual field ahead or on the horizon) is where the participants' attention is directed (360°). Visuals are on the anterior horizon and are emitted from a central point of view at the center of the visual horizon. The participant can move their head around and continue to experience the visual effects that appear before them, but the content is not dynamic in real time nor does the content react to the participants' movements. Color and sound changes mark subtle content transitions. Additional elements such as specific frequencies of binaural audio are incorporated to effect transitions.

**Figure 2.** Representative images of virtual reality experience (courtesy of Harmony Media).



## Surveys

See Multimedia Appendix 1 for full surveys. To measure the degree to which situations in one's life are appraised as stressful, with a time recall of 1 month, the PSS-10, one of the most widely used psychological instruments for measuring perception of stress [31-34], was scored and summarized and then compared to sex, age, and race norms [34]. The primary outcome was (presurvey vs postsurvey) change in STAI-S scale (Δ-STAI-S). It is meant to measure, via self-report, the presence and severity of current symptoms of anxiety and a generalized propensity for anxiety. The State Anxiety Sub-Scale asks respondents how they feel "right now" using items that measure subjective feelings of apprehension, tension, nervousness, worry, and activation or arousal of the autonomic nervous system using a 4-point Likert scale. Scores for each scale range from 20 to 80, with higher scores indicating greater stress or anxiety (clinically significant symptoms suggested with a score of ≥39 - 40).

Several studies have used the State Anxiety Scale to measure change immediately following situational stress-inducing or stress-reducing interventions, and they show that the instrument is sufficiently responsive to capture short-term changes [35-40]. To explore immersive tendencies as a predictor of responsiveness to VR, the ITQ, which is used to measure differences in individual tendencies to experience presence or immersion, was used, while the SSQ was included to evaluate the safety and tolerability of VR by measuring the frequency of acute VR discomfort (eg, headache, vertigo, nausea, and eye strain) resulting from sensory mismatch between the visual and vestibular systems called "cybersickness." Overall rating (0 being the worst to 10 being the best) and ITQ responses were scored and analyzed using summary statistics, while the SSQ scores were interpreted using 5 standardized categories [53].

## Physiologic Data

HRV reflects complex neurocardiac regulatory interactions, including the balance of inputs from the autonomic nervous system. For HR and HRV variables, raw R-R intervals for each participant were exported from EliteHRV to Premium Kubios HRV Analysis Software (version 3.4.3; Kubios Oy). Preprocessing of the data used the automatic beat correction algorithm built into the Premium Kubios packaging [54]. The 30-minute (during VR) recording was then broken into six 5-minute periods, with each HRV variable for the period extracted for analysis. This allowed for the creation of longitudinal models across 8 time points, given that the HRV variables examined are valid with a recording period of 5 minutes [55-57]. These time points are denoted as (1) pre-experience (−5 to 0 minutes), during experience: (2) 0 to 5 minutes, (3) 6 to 10 minutes, (4) 11 to 15 minutes, (5) 16 to 20 minutes, (6) 21 to 25 minutes, (7) 26 to 30 minutes, and (8) postexperience (31 to 35 minutes). Longitudinal models provided greater granularity of analysis with potential pattern recognition across the cohort across time points.

Two HRV time-domain variables were analyzed. Root-mean-square of successive differences (RMSSD) is considered the most relevant and accurate measure of autonomic nervous system activity and specifically estimates vagal-mediated changes in HRV [56]. Standard deviation of the interbeat interval of normal sinus beats (SDNN) is another overall estimate of HRV, and in short-term recordings, mainly reflects parasympathetic input predominantly from slow-paced breathing [56]. HRV frequency domain results included the following variables (analyzed using the Fast Fourier Transformation method): (1) low frequency (LF) to high frequency (HF) ratio (LF:HF): while not exact, this measure can indicate the balance of sympathetic versus parasympathetic tone. Decreases in this score reflect increased parasympathetic

(or decreased sympathetic) tone. The following absolute and relative power of the HRV frequency domains were also included: (2) very low frequency band (0.0033 - 0.04 Hz): it provides insight into parasympathetic activity, (3) LF band (0.04 - 0.15 Hz): this band mainly reflects baroreceptor activity during resting conditions and can reflect both sympathetic and parasympathetic tone, and (4) HF band (0.15 - 0.40 Hz): this band reflects vagal tone and is referred to as the respiratory band because it is associated with variations in HR related to the respiratory cycle. Lower HF power is correlated with stress, panic, anxiety, or worry] [55-57].

GSR is another parameter that detects changes in autonomic activity through sweat gland function. Heightened psychological and physiological arousal (eg, stress) increases sympathetic activation leading to increased skin conductance detected by the GSR sensor [46].

### Statistical Analysis

Descriptive statistics of the cohort's demographics (including experience with VR previously) were summarized along with cardiac conditions or risk factors gathered from the electronic medical record; quantitative variables were formatted and summarized using Microsoft Excel (version 16.75) with Real Statistics add-on and further analyzed with RStudio (version 2023.06.1+524; Copyright 2022 by Posit Software, PBC) using either 2-sample paired 2-tailed *t* tests or the nonparametric equivalent, Wilcoxon signed rank sum test, along with 95% CIs. A significance level of $\alpha<.05$ was used. Subsequent BP, HR, HRV, and GSR data were compared to pre-experience (baseline) measurements to look for statistically significant changes across time points. A multivariable linear regression model was used to evaluate for potential survey variable predictors of the change in stress as measured by our primary outcome variable, Δ-STAI-S scale score. The effect size (*r*) for the Wilcoxon paired signed rank test was computed for Δ-STAI-S by dividing the *z*-value by the square root of the sample size (corresponding to the total number of pairs, n=20). The *r* value varies from 0 to 1, and we used the following interpretation cutoffs commonly published in the literature: 0.1 to <0.3=small effect, 0.3 to <0.5=moderate effect, and ≥0.5=large effect. To further contextualize our findings, we identified the threshold for discriminating a minimal clinically important difference as 5 units on the STAI-S scale. This is a conservative estimate based on the widely accepted half a SD (½ SD) rule of thumb for interpreting changes in health-related quality of life instruments [58-60]. For the STAI-S scale with 20 items each with a 1 to 4 scale, this creates a range from 20 to 80 or 60 points (1 SD=10 units and ½ SD=5 units). For complete physiologic and survey datasets, please see Multimedia Appendix 2.

### Qualitative Analysis

Interviews were analyzed using reflexive thematic analysis, a qualitative research method in which the researcher not only identifies themes but also reflects during the process on how their own interpretations and biases might influence the analysis [61]. A medical student conducted the interviews and analysis. Using a semistructured interview guide (Multimedia Appendix 1) with open-ended questions while maintaining a neutral and welcoming tone throughout the study visit helped to ensure that the data collected reflected the participant's honest feedback. The same open-minded approach was taken in the analysis, using an overall social constructivist theoretical framework [62]. An inductive approach was used, and analysis sought to reflect all the participants' actual responses rather than only code and theme development being directed by existing concepts, theories, and ideas. This allowed for a broader understanding of how the participants' personal preferences and feedback were informed by previous experiences, knowledge, and abilities. Familiarization with the data was done by reading through the transcription multiple times with annotation for potential areas of interest or repetition of topics or concepts. A total of 11 initial codes were then generated with examples of text supporting each code being labeled and sorted under initial themes. A "critical friend," a volunteer researcher with previous qualitative analysis experience, was consulted after coding to help talk through and clarify ideas while providing impartial feedback. They were given access to original interview transcripts so that they could make their own coding decisions. This helped to ensure the quality of analysis by challenging the primary analyst to adopt a more independent stance toward the research and ensure the analysis was coherent. The codes and supporting examples were then reviewed; themes or sub-themes were constructed and refined.

### Mixed Methods Analysis

Roughly equal weight was given to both qualitative and quantitative data while exploring the same question. Results from each data stream were compared to see if they reached the same conclusions through triangulation or in a complementary way, with qualitative data providing greater depth of understanding to quantitative results and vice versa [63].

## Results

### Recruitment and Study Population

In total, 48 individuals were referred as potential participants (provider-referred: n=37, self-referred through flyer: n=7, ClinicalTrials.gov or other: n=4). Of these, 42 were screened for eligibility via a phone call with a study staff member, 5 did not meet inclusion criteria, 17 declined participation, and 20 were enrolled or completed participation. The remaining 6 referred individuals were called by study staff but were unable to be reached. A total of 20 participants were included for all demographics or survey outcomes, while 18 included were for physiologic variables. Technical difficulties were encountered during the study visit for participant 17, leading to a lack of physiologic data. Data from participant 10 were consistently found to be a statistical outlier, likely due to the fact that they were a heart transplant recipient. Participant demographics, cardiac conditions or risk factors, and other comorbidities are shown in Table 1. The mean age of participants was 66 (SD 15; range 27-80) years. There were 9 female and 11 male participants. The most common stress relief activity was exercise, followed by deep breathing and then meditation (Multimedia Appendix 1).

**Table .** Participant characteristics.

| Patient characteristic | | Values (N=20) |
|---|---|---|
| Age (years) | | |
| | Mean (SD) | 66 (15) |
| | Range | 27-80 |
| Sex (assigned at birth), n (%) | | |
| | Male | 11 (55) |
| | Female | 9 (45) |
| Race, n (%) | | |
| | African American or Black | 0 (0) |
| | Asian | 3 (15) |
| | White | 14 (70) |
| | Other races | 3 (15) |
| Hispanic, n (%) | | 3 (15) |
| Education, n (%) | | |
| | Some college | 2 (10) |
| | College degree | 11 (55) |
| | Advanced graduate degree | 7 (35) |
| Household income (US $), n (%) | | |
| | <$50,000 | 3 (15) |
| | $50,001-$100,000 | 6 (30) |
| | $100,001-$200,000 | 3 (15) |
| | >$200,000 | 6 (30) |
| | Prefer not to say | 2 (10) |
| Employment status, n (%) | | |
| | Part-time | 1 (5) |
| | Full-time | 6 (30) |
| | Retired | 9 (45) |
| | On disability | 1 (5) |
| | Unemployed | 1 (5) |
| | Homemaker | 2 (10) |
| Insurance status, n (%) | | |
| | Current or former employer | 11 (55) |
| | Direct from company | 1 (5) |
| | Medicare | 6 (30) |
| | Medicaid | 1 (5) |
| | Other | 1 (5) |
| Relationship status, n (%) | | |
| | Never married | 4 (20) |
| | Married | 11 (55) |
| | Divorced | 1 (5) |
| | Separated | 1 (5) |
| | Widowed | 3 (15) |
| Prior experience with virtual reality, n (%) | | |

| Patient characteristic | | Values (N=20) |
|---|---|---|
| | No experience | 11 (55) |
| | A little bit of experience | 5 (25) |
| | Some experience | 2 (10) |
| | Quite a bit of experience | 2 (10) |
| Hyperlipidemia, n (%) | | 14 (70) |
| Hypertension, n (%) | | 11 (55) |
| Coronary artery disease, n (%) | | 8 (40) |
| Overweight (BMI 25 - 30), n (%) | | 8 (40) |
| Valvular disease or replacement, n (%) | | 7 (35) |
| Former smoker, n (%) | | 7 (35) |
| ≥1 Myocardial infarction or multivessel CABG[a], n (%) | | 5 (25) |
| History of atrial fibrillation, n (%) | | 5 (25) |
| Obesity (BMI≥30), n (%) | | 5 (25) |
| Diabetes mellitus type 2, n (%) | | 4 (40) |
| Diastolic heart failure, n (%) | | 3 (15) |
| Currently in cardiac rehabilitation, n (%) | | 3 (15) |
| Previous cardiac rehabilitation, n (%) | | 3 (15) |
| Heart transplant, n (%) | | 1 (5) |
| Anxiety, n (%) | | 9 (45) |
| Depression, n (%) | | 9 (45) |
| Chronic pain, n (%) | | 7 (35) |
| Insomnia, n (%) | | 4 (20) |
| Migraine, n (%) | | 3 (15) |

[a]CABG: coronary artery bypass graft.

## Survey Outcomes

### *PSS-10, STAI-S, and ITQ*

Baseline PSS-10 results showed a mean score of 19 (SD 6; range 4-32). Comparing each PSS-10 score to the sex, age, and race-related norm categories provided in the PSS-10 item inventory revealed that 18 of 20 participants had greater than average stress [31]. Pre-STAI-S scale also suggested elevated current stress with a mean of 35 (SD 12), median of 31 (IQR 28-38), with a range of 21 to 71 on a possible scale of 20 - 80. In total, 1 had “no anxiety” (20-21), 14 had “low anxiety” (22-37), 2 had “moderate anxiety” (38-44), and 3 had “high anxiety” levels (45-80) [40].

The primary outcome, Δ-STAI-S scale, showed a median decrease of 7 (median 31, IQR 28-38 vs median 24, IQR 21-29; $P$<.001) and an average decrease of 8.5 (SD 9.8) or 21%. This corresponded to a large effect size ($r$=0.77). Both the median and average change were larger (ie, more negative) than the threshold of clinical significance set at −5 units. In total, 12 of 20 (60%) participants reached this threshold. A total of 1 of 20 (5%) participants had an increase equal to the threshold (ie,+5 units).

The mean for the cohort on the ITQ was 59 (SD 12.6) with a range of 47 to 90 on a maximum scale of 18 to 126. The majority of participants reported negligible (n=11) or minimal (n=1) “cybersickness,” while 5 had “significant” symptoms, and 3 fell into the “bad” range [53]. Of note, those who did report symptoms rated each as mild or moderate. None were rated as severe. Fatigue was the most commonly reported symptom. The mean rating of the VR experience overall was 9 of 10 (SD 2; median 9, IQR 8-10; mode 10) with one outlier who gave a rating of 3. In total, 11 individuals had no previous experience with VR, 2 had a little bit, 1 had some, and 2 had quite a bit.

Finally, multiple regression analysis was conducted to examine the relationship between the variable: Δ-STAI-S scale and various potential predictors gathered through the surveys including STAI-S pre-experience score, PSS score, SSQ score, ITQ score, overall VR rating, age, and sex. The initial model demonstrated good fit and accounted for 86% of the variability in Δ-STAI-S (multiple $R$=0.96; adjusted $R^2$=0.86; ANOVA of regression: $P$<.001). However, only the STAI-S scale pre-experience score ($P$<.001) and overall rating ($P$=.04) were significantly correlated with Δ-STAI-S. After excluding the nonsignificant variables, the model showed that both variables had negative coefficients that were significantly correlated with Δ-STAI-S, indicating that those with higher scores on these

variables tended to have a greater decrease in STAI-S scale. The model with the final 2 predictors produced multiple $R$=0.95, adjusted $R^2$=0.90, ANOVA of regression: $P$<.001 (Multimedia Appendix 1).

### *Physiologic Outcomes: BP, HR, and GSR*

There was no significant change in systolic or diastolic BP among participants (Table 2). Pre- versus postexperience HR recordings demonstrated a statistically significant decreased mean HR of nearly 6 beats per minute. Upon examination of longitudinal models with spaghetti plots, it was observed that this drop tended to occur between time point 1 (pre-experience) and time point 2 (0 to 5 minutes), followed by a plateau (Figure 3). Note that findings remained significant when adjusted for multiple statistical comparisons using the Bonferroni correction as well as Holm and Hochberg tests. The change in GSR conductance and area under the curve (before VR vs after VR) did not demonstrate a significant change (Table 2). GSR was only analyzed before and after the VR experience because of technical issues that affected data collection for 5 of 20 participants.

**Table .** Change in State Trait Anxiety Inventory-State (STAI-S) scale, blood pressure (BP), and galvanic skin response (GSR) (N=20).

| Variable | Pre-VR[a] | | Post-VR | | Δ (post-pre)[b] | | 95% CI | *P* value |
|---|---|---|---|---|---|---|---|---|
| | Mean (SD) | Median (IQR) | Mean (SD) | Median (IQR) | Mean (SD) | Median (IQR) | | |
| STAI-S (points) | 35 (12) | 31 (28-38) | 26 (6) | 24 (21-29) | −8.5 (9.8) | −7 (−11.5 to −3) | −11.5 to −4 | .001[c] |
| Systolic BP (mm Hg) | 128 (14) | 127 (122-133) | 129 (18) | 128 (115-138) | 1 (13.8) | 1 (−9 to 6.5) | −5.5 to 7.5 | .75 |
| Diastolic BP (mm Hg) | 79 (9) | 78 (71-85) | 80 (10) | 79 (75-89) | 1 (7) | 1 (−3 to 4) | −2.3 to 4.2 | .60 |
| Heart rate (beats per minute) | 73 (8) | 72 (67-78) | 67 (6) | 65 (63-70) | −6 (5) | −7 (−8 to −3) | −8.4 to −3.5 | <.001 |
| GSR (microsiemens) | 0.74 (0.89) | 0.48 (0.28-0.79) | 0.70 (0.57) | 0.61 (0.32-0.82) | −0.036 (0.42) | 0.11 (−0.17 to −0.3) | −0.2 to 0.2 | .50[c] |
| GSR AUC[d] (microsiemens*$10^{-1}$ second) | 1979 (2518) | 1308 (836-2162) | 1933 (1701) | 1726 (854-2191) | −46 (1183) | 317 (−307 to 504) | −256 to 448 | .41[c] |

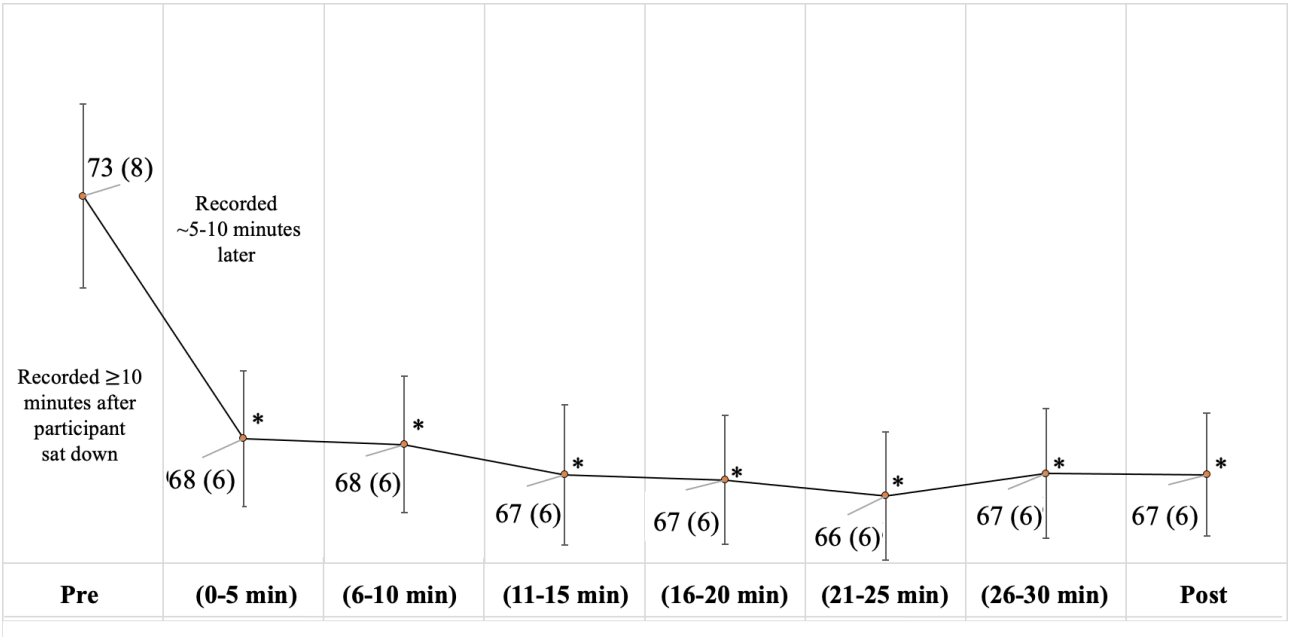
[a]VR: virtual reality.

[b]Δ: change or difference.

[c]Indicates Wilcoxon signed rank sum test (otherwise 2-sample paired *t* test used).

[d]AUC: area under the curve.

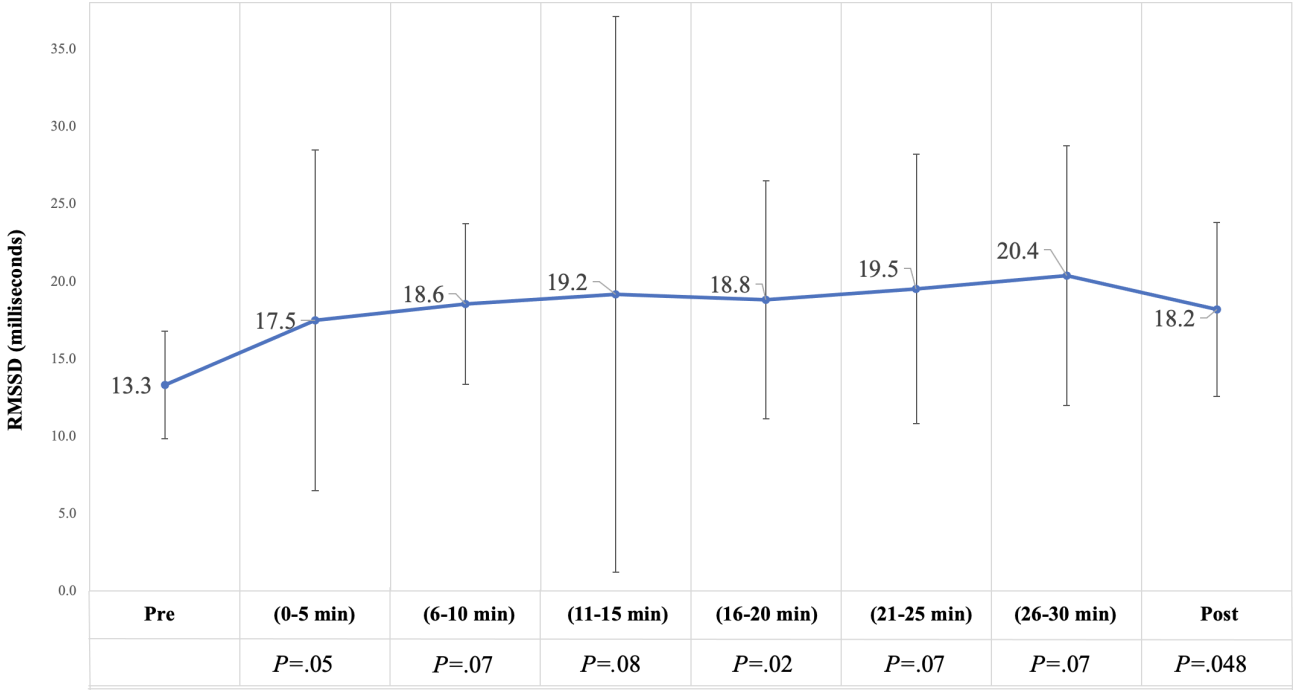
**Figure 3.** Change in average heart rate (beats per minute). Error bars represent SE. Post: recorded 5 minutes after experience; Pre: recorded 5-10 minutes before experience. *$P$<.001.



### *Physiologic Outcomes: Time-Domain HRV Parameters*

The median RMSSD (milliseconds) for the cohort was increased at all time points relative to baseline (highest at time point 26 - 30 minutes) but was significantly increased from baseline only at 16 - 20 minutes and after the experience (Figure 4). Median SDNN (milliseconds) also tended to increase across time, though the change from baseline was only statistically significant after the experience (Multimedia Appendix 1).

**Figure 4.** Change in median RMSSD (milliseconds). Error bars represent SE. Post: recorded 5 minutes after experience; Pre: recorded 5-10 minutes before experience; RMSSD: root-mean-square of successive differences.




### *Physiologic Outcomes: Frequency-Domain HRV Parameters*

In terms of relative power, there was a significant decrease in very low frequency power (relative to pre-experience or baseline) at 3 time points, associated with a reciprocal increase in HF, also significant at 3 time points (Figure 5). The LF:HF ratio was decreased relative to baseline at all time points, reaching a nadir in the middle of the experience and gradually increasing again during the last 3 time points (Figure 6). All median values were greater than 1, indicating that at least half of the participants had greater power from LF than HF. Note that none of the HRV findings were still significant after adjusting for multiple statistical comparisons using the Bonferroni correction and Holm and Hochberg tests.

**Figure 5.** Change in relative frequency band power. HF: high frequency; LF: low frequency; Post: recorded 5 minutes after experience; Pre: recorded 5-10 minutes before experience; VLF: very low frequency. *$P$<.05 (see Multimedia Appendix 1 for specific values).

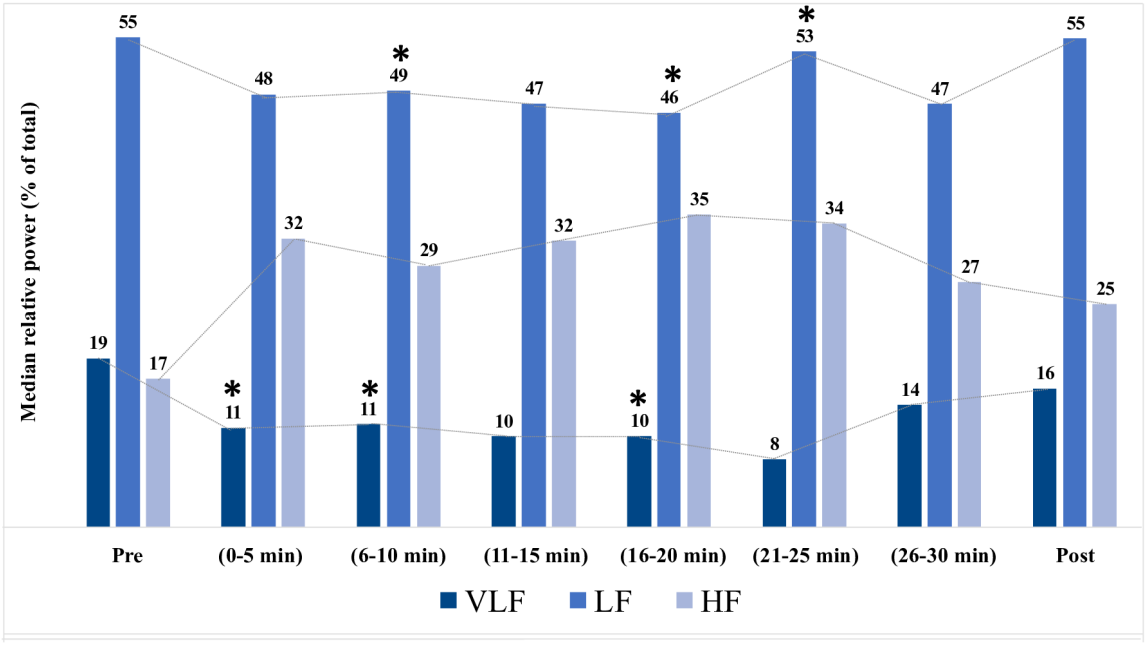
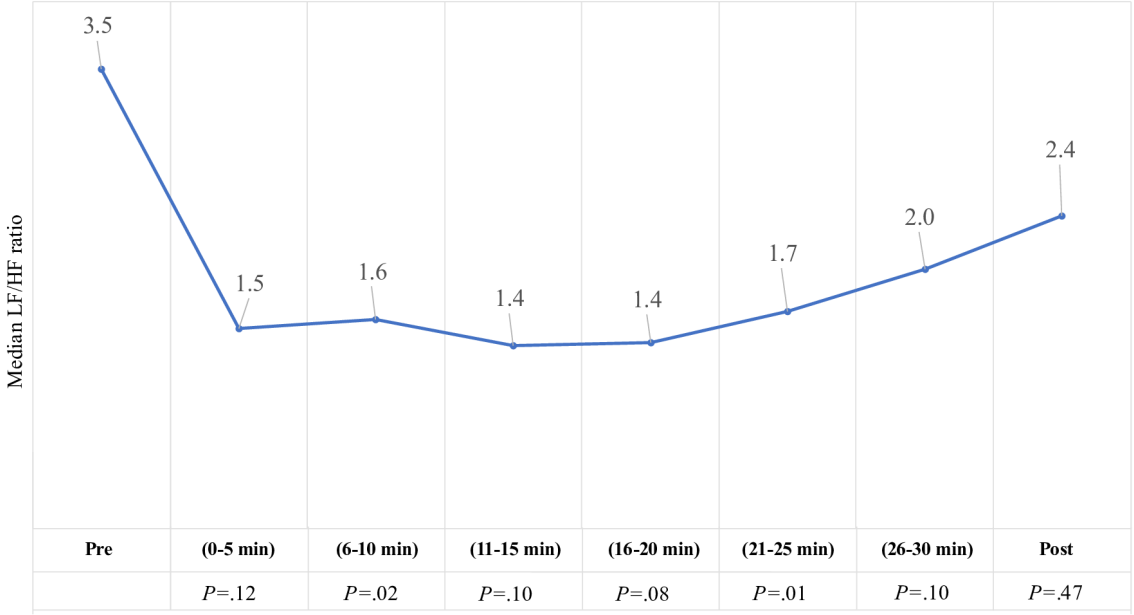



**Figure 6.** Change in median LF:HF ratio. HF: high frequency; LF: low frequency; Post: recorded 5 minutes after experience; Pre: recorded 5-10 minutes before experience.

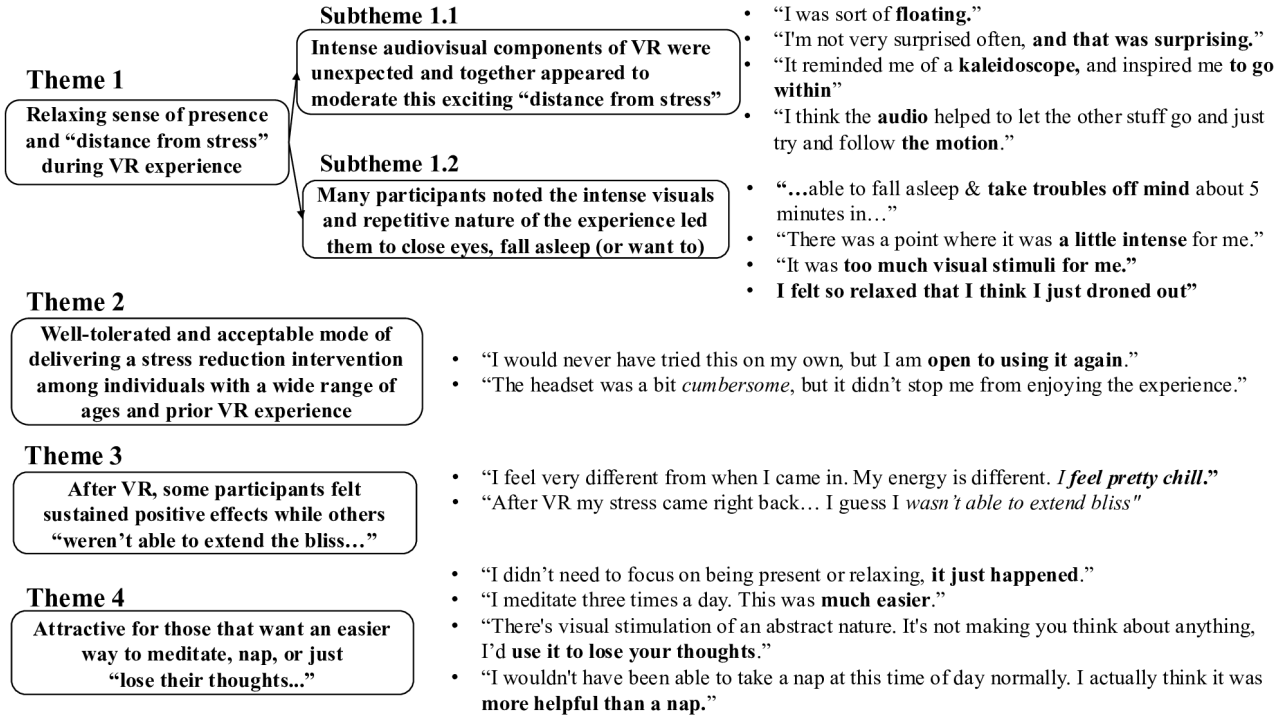



## Qualitative Results

### *Overview*

In total, 6 potential themes were constructed. Upon reviewing and defining the themes, 2 themes did not have sufficient evidence to constitute substantive themes and rather reflected subthemes of theme 1. In the interviews, participants described their stress and anxiety as external and out of their control rather than describing an internal locus of control. For example, some described worry about family members, painful divorce, or daily stresses piling on from career, traffic, or their roof leaking. Others expressed stress more directly related to their cardiac conditions or adverse cardiovascular events. Figure 7 summarizes the main themes and subthemes that emerged regarding the VR experience.

**Figure 7.** Qualitative interview results using inductive thematic analysis: 4 themes and 2 subthemes. VR: virtual reality.




### *Theme 1*

During VR, participants experienced a relaxing sense of presence and "distance from stress." Participants generally reported losing track of all sense of time or time passing more quickly while they had a sense of "sort of floating." Others completely forgot about or felt removed from usual stress, with one person stating, "I wasn't able to focus on other thoughts really" and another said, "I could lose myself in the patterns and everything and kind of just ... I don't know, this is something different." Qualitative analysis also provided insight that the intense audiovisual components of VR were unexpected and appeared to moderate this "distance from stress" (subtheme 1.1). Participants described that the combination of audio and visuals synergistically created an immersive experience that created an almost "indescribable factor." Many participants organically offered that they were pleasantly surprised by what they experienced. Some mentioned that the red and yellow colors drew them in, while others focused on the starry sky transporting them back to being a child star gazing. Others found that the sacred geometry reminded them of a "kaleidoscope" inspiring them to "go within" or that the "motion around the geometry with the concentric circles [paired with] the audio helped to let the other stuff go." Many participants were most surprised by the nonverbal audio, with one person noting, "I wouldn't have listened to that on my own, but surprisingly I liked it." Some described that they would have expected to experience something different. For one person, these intense visuals were "off-putting" with "too much going on," and envisioning the experience would have something more relaxing like "someone walking slowly on a path."

The next subtheme (1.2) was the intense visual experience that led participants to close their eyes (or at least want to). On one end of the spectrum, some participants felt so relaxed that they felt as if they fell asleep after about 5 minutes. On the other end of the spectrum was the patient mentioned earlier who said it was "difficult to keep my eyes open" due to eye strain. Another noted, "there was a point where it was a little intense for me."

### *Theme 2*

The VR experience was well-tolerated overall and was an acceptable delivery mode for this stress reduction intervention across a wide range of ages. A few participants noted that the headset was slightly cumbersome or that their head felt slightly turned or not centered, "I loved the experience, it was very relaxing, it would have been perfect had there been a better fit with the headset/headphones." However, most participants stated that the headset was not uncomfortable and did not hinder their enjoyment of the content. There were few other physical sensations discussed during the interview aside from one participant describing a transient moment of "stomach dropping" and one participant who experienced eye strain. Universally, participants were open to trying VR again, including the one with eye strain (though with different content).

### *Theme 3*

After taking off the headset, some participants had sustained positive effects, while others were "not able to extend the bliss." On one end of the spectrum, one participant noted that his busy mind and daily stresses came flooding back to him when he took off the headset and especially once he was handed the postexperience survey. Another noted, "I mean, I do feel relaxed. I do feel present. I do feel energized and clear, much more clear than when I came in the room. But now there's also that old [stress] too, of like yeah, I remember that." Others had a change in mood and energy that was visible and palpable to the interviewer, with one participant feeling more confident, "I just feel more sure of myself. And I'm able to push all my worries about my daughter, more away." One person reflected, "my

energy is different now ... under stressful times, you're thinking about [many things], but then when I did this, I was just like ... now I feel pretty chill." While poststudy visit follow-up was not part of the protocol, one participant reached out the next day just to the study team to say that they had "felt a sense of calm and peace that last well into the evening ..."

### *Theme 4*

Participants felt that this VR experience may be helpful for those who want an easier way to meditate, nap, or just "lose their thoughts" for a bit. Multiple participants referred to the "abstract nature of the visual stimulation" in particular or losing themselves in the patterns. A few participants compared the VR experience to other stress reduction methods, such as meditation, progressive muscle relaxation, or even taking a nap. They proposed that it could be used "as an interlude for daily stress" and would recommend it to others as "a good tool to mellow out in the moment." While the state of mind and feelings seemed to be described as similar to those induced by these techniques, they described that "[VR] was much easier" and for one they thought it could even "replace my regular meditation practice." This final theme also underscores the participants' experience of this VR intervention and future therapeutic value.

### Mixed Methods Analysis

The VR intervention by mixed methods analysis had a positive effect on stress reduction based on (1) the statistically (median STAI-S $\Delta$=–7*u*; *P*<.001) and clinically significant (minimal clinically important difference=STAI-S scale $\Delta$5*u* or one half of a SD for this health-related quality of life scale) [59] decrease in STAI-S scale scores, (2) the statistically significant decrease in HR (−6 beats per minute; *P*<.001), and (3) qualitative interview feedback describing that the intervention induced a relaxing sense of "distance from stress" that was associated with a change in energy and feelings similar to meditation or a flow state. However, the remaining physiologic outcomes BP, GSR, and HRV were not significantly changed, and thus moderated the magnitude of this positive effect. Overall, survey data, physiologic data, and qualitative interviews support the conclusion that the VR intervention is safe and tolerable. Interview feedback provided further depth, with most participants describing an enjoyable experience that they were open to doing again.

## Discussion

### Overview

Our study found that a sample of patients with CVD or risk of CVD had above-average stress. A statistically and clinically significant decrease in subjective perception of stress was observed after participants experienced a novel VR application. HR significantly decreased, while BP and GSR remained the same. Changes in HRV parameters were consistent with increased vagal tone over time but were only statistically significant at certain time points (before multiple comparisons adjustment). Overall, the VR intervention was a safe and feasible stress reduction method.

### Principal Findings

Our pilot study of a VR experience for patients with or at risk for CVD demonstrated several interesting findings. First, this sample of patients had higher than average stress based on preintervention stress scores, with nearly half having EMR evidence of anxiety or depression, which is consistent with previous epidemiological data of patients with coronary artery disease and those who have experienced severe cardiac events [1,3-7]. This finding also suggests a need to address mental stress in patients with or at risk for CVD. The STAI-S scale is widely validated with high internal consistency (0.86 - 0.95) and reliability (0.65 - 0.75) coefficients [40], while PSS-10 was also chosen for its validity and reliability [32]. Higher PSS-10 scores have been associated with failure to quit smoking, worse control of blood sugar, greater vulnerability to depressive symptoms, and more frequently having the common cold [31]. Qualitative analysis provided additional scope to the finding of elevated average stress in our cohort (Figure 7).

Second, there was a statistically (median $\Delta$=–7*u*; *P*<.001) and clinically significant (minimal clinically important difference=$\Delta$5*u* or one half of a SD for this health-related quality of life scale) [59] decrease in subjective perception of stress measured by survey data that matched theme 1 and was overall supported by qualitative feedback. Previous research suggests that VR exerts its effect by temporarily silencing the DMN and can lead to the state of mind opposite to that produced by the DMN, coined "flow" [64]. Often described as "being in the zone," the 4 attributes of flow include selflessness, timelessness, effortlessness, and richness [65]. It is thought that flow requires "transient hypofrontality" or a temporarily suppressed prefrontal cortex that allows the self-critical mind to quiet and connect neural areas that do not normally communicate [66]. This "lateral thinking" [18] can create the opportunity for new insights and a profound sense of calm.

VR experts propose the following mechanism by which VR can induce flow within minutes: VR is completely immersive, so that it can capture the user's full attention like no other audiovisual medium. By removing distractions and offering a novel, information-rich experience, it can interrupt the noisy DMN, change brainwaves from $\beta$ to $\alpha$, and increase neurohormonal signaling through dopamine, oxytocin, and serotonin that has the effect of creating an aura of tranquility that can conjure flow [18]. VR was shown to produce a nearly identical state of consciousness as psilocybin, a known flow-inducing compound, using a VR program called the Hallucination Machine that radically altered visual perceptions [20]. While psychedelics have been an area of booming interest for the treatment of severe anxiety, phobias, depression, autism, and posttraumatic stress disorder, these findings demonstrate exactly why VR continues to expand in the area of psychiatry and can be applied generally to stress reduction.

The physiologic data did not uniformly converge with these subjective findings. On the one hand, HR significantly decreased, and HRV parameters showed significant changes consistent with increased parasympathetic state at certain time points. Pre- versus post-HR recordings demonstrated a statistically significant decreased mean HR of nearly 6 beats

per minute. Longitudinal plots showed that the initial drop tended to occur between the first 2 time points and then persisted. The protocol attempted to allow time for participants to reach resting HR before the first recording (completing consent, pre-experience survey, and connecting chest strap and GSR sensors while in a seated position). For most healthy adults, HR stabilizes after 4 minutes of inactivity [67]. So, it is possible that this represents a true effect related to VR. However, if our sample of participants required longer for cardiovascular activity to return to the resting state due to their health state or there were other factors that interfered (such as mental stress provoked by completing the presurvey), then it is plausible that the observed results are due to the passage of time. This cannot be deciphered without having a placebo-controlled trial.

RMSSD (mainstream apps such as Elite HRV and Whoop report this as an "HRV score," with higher scores generally indicating better "readiness" for the body to adapt and perform), SDNN, relative power of HF band, and LF:HF ratio showed changes consistent with a more relaxed state between the first 2 time points. HRV findings were most consistent with increased vagal tone. The regulatory mechanisms at play for "short-term" HRV measurements (5-minute recordings) include changes in HR driven by changes in respiration (respiratory sinus arrhythmia), the baroreceptor reflex (negative-feedback regulation of BP), and rhythmic changes in vascular tone [56]. A period of HRV monitoring can be studied using time-domain parameters, which look at aspects of the variability in the interbeat interval, and frequency-domain parameters, which look at the distribution of power generated by each beat (by breaking the total power into 4 frequency bands) [56,57]. HRV provides a measure of how well the cardiovascular system can rapidly adjust in response to sudden physical and psychological disruptions to homeostasis. The Fast Fourier Transformation method was preferred over the autoregressive HRV analysis method because several studies have found that the autoregressive method did not produce reliable results in some patients with diabetes or hypertension [68], which we expected to be prevalent in our population.

On the other hand, BP, GSR, and at several time points, HRV parameters showed no significant change. Further, when adjusting for multiple statistical comparisons, HR changes remained significant, while HRV changes were no longer significant, indicating that these findings may be an artifact or type I error. It has been repeatedly shown that breathing exercises can cause a modest but significant decrease in HR (−2.41 beats per minute; $P$=.03) and BP (systolic: −7.06; $P$<.01 and diastolic: −3.43; $P$<.01) [69]. If there was a change in breathing pattern or vagal tone, consistent with the observed HRV trends, it could partially explain the decreased HR, but we would have also expected BP to fall. Note, however, that BP was not measured longitudinally to minimize patient disturbance. Similarly, we hypothesized that there would be a significant reduction in GSR since some studies have shown that during meditation and after consistent meditation, participants had a statistically significant reduction in GSR compared to controls [47-49]. All of these physiologic measures are subject to influence from diurnal fluctuations, positioning, recency of physical activity, caffeine intake, and sleep [56,70]. These conditions were not standardized in our study.

It is also important to note that this was a pilot study without a control; it aimed to gain initial insights into the potential effects of the VR intervention before embarking on a full-scale randomized controlled trial. The small sample size and lack of a control group limit our ability to conclude that these observed results were due to the VR intervention, specifically. There is currently a need for studies with a larger, more diverse population and control group to be able to adequately assess the effects of this VR intervention. This study can inform such studies. Although there were mixed effects in changing physiologic measures in this study, there is other evidence that VR can lower the startle response and stress hormone levels for 12 months after treatment [28] and normalize brain function confirmed through functional magnetic resonance imaging—suggesting that VR can clearly impact the mind but also the body [29].

Importantly, quantitative and qualitative analyses showed that the VR intervention was not only a tolerable and feasible stress reduction method, but it was also enjoyable across a wide range of (1) ages and (2) prior experience with VR. The SSQ produced reassuring results that are consistent with current estimates of VR cybersickness. Approximately 60% - 95% of users experience some level of cybersickness, with 5% - 15% ending their experience prematurely due to symptom severity. The prevalence and significance of cybersickness continue to fall with technical improvements in both hardware and software [71,72]. Our intervention uses slow-moving visuals that build gradually, which hopefully makes it less jarring than highly kinetic scenes. However, the 30-minute duration may contribute to the development of cybersickness. One of the most commonly reported cybersickness symptoms was fatigue. Resting in the semirecumbent position may have contributed to this. A total of 16 of 20 (80%) participants had little or no VR experience. Universally, participants were open to trying VR again, including the one with eye strain (just with different content). According to the ITQ, the cohort varied in their likelihood to become involved or engrossed in different environments and generally tended to be prone to "moderate immersion." While this tendency supports the idea that the effectiveness of VR is moderated by a sense of presence in the experience, there was not a statistically significant relationship between ITQ and the outcome measure.

To further explore the potential moderators of the observed changes in the STAI-S scale, we performed regression analysis. Multilinear regression findings highlight that those with higher baseline acute anxiety scores tended to have larger decreases in acute anxiety. Higher overall ratings also predicted greater Δ-STAI-S scores, possibly because they valued the experience due to its anxiety-relieving effects. Per our regression model, baseline acute anxiety and overall VR rating explained 90% of the variation in STAI-S scale change. Interestingly, age, sex, chronic stress level over the past month, SSQ, and ITQ scores were not correlated with the primary outcome, nor did they have any explanatory value for variation in STAI-S scale change. Simple linear regression demonstrated weak to moderate negative correlations between age and PSS ($r$=−0.28; $P$=.23)

and STAI-S scale prescores ($r$=−0.48; $P$=.03), suggesting a trend that older participants tended to have lower current and chronic stress levels. There was a moderate negative correlation between age and magnitude of decrease in STAI-S scale ($r$=−0.46; $P$=.05), suggesting that younger participants with more baseline stress experienced greater stress reduction.

## Comparison to Prior Work

To provide context for our primary outcome, Δ-STAI-S, we compared our findings (mean −9, SD 10; median −7) with previous studies using the same outcome variable to measure stress and anxiety reduction from a VR intervention (Table 3) [73-77]. While there are several differences between these studies, they provide support that the VR content under investigation may be capable of producing similar or stronger effects compared to other VR interventions. All 6 studies demonstrated statistically and clinically significant ($P$<.05; Δ≤–5 units) change from baseline, adding to the growing body of evidence for VR as a powerful stress reduction method. Of note, Baytar and Bollucuo Lu [74] incorporated hemodynamic measures of change and saw a similar initial decrease in HR of 6 beats per minute, which was statistically significant, with an additional 2 beats per minute decrease subsequently. Only 2 studies were found that examined the effect of VR on STAI scores in a group of patients with CVD; however, both used the 6-item short form of the STAI, limiting comparison with this study [78,79]. These studies, like many others in the literature, used VR for periprocedural anxiety.

**Table .** Comparison of this study to previous studies using virtual reality (VR) and change in State-Trait Anxiety Inventory-State Scale as an outcome for stress or anxiety reduction.

| Study | Description | Change[a] | *P* value |
|---|---|---|---|
| Makaroff et al | This study | • −9 (SD 10) (median change −7) | • <.001[b] |
| Kim et al (2021) [73] | Randomized crossover design measuring anxiety of 74 healthy adults with high stress after experiencing intentionally stressful VR followed by either BF[c] or VR relaxation | • −6 (SD 10) (Stress-VR)<br>• −6 (SD 8) (Stress-BF) | • VR versus BF: 0.39<br>• VR:<.001, BF:<.001 |
| Baytar and Bollucuo Lu (2021) [74] | 40 patients undergoing septorhinoplasty with 15 minutes of preoperative 360-degree VR content with nature scenes and “meditation music” | • −7 (median change −7) | • <.001[b] |
| Niki et al (2021) [75] | 10 participants in a Japanese nursing home with two 10-minute immersive VR slideshows | • −9 (after first VR)<br>• −13 (after second VR) | • <.001 |
| Brown and Foronda (2020) [76] | 7 patients undergoing outpatient surgeries viewing AppliedVR modules during surgery | • −13 | • .03 |
| Karaman and Taşdemir (2021) [77] | 30 female patients using VR during fine-needle aspiration breast biopsy versus 30female patients undergoing standard protocol | • −12 (experiment)<br>• −7 (control) | • <.001 |

[a]SD and median were not available for several studies.

[b]Indicates Wilcoxon signed rank sum test, otherwise 2-tailed *t* test used.

[c]BF: biofeedback.

## Strengths and Limitations

This clinical investigation of a novel VR intervention designed for stress reduction in patients with CVD or at cardiovascular risk has several strengths. The pilot examined both physiologic and subjective patient data. Our inclusion of longitudinal HRV analysis helps to better explore and measure how VR may influence complex neuro-cardiac interactions that result in less stressful states. The mixed methods design and inclusion of interviews or qualitative data analysis add richness and depth to the conclusions that could be drawn from a purely quantitative or qualitative analytical approach.

Our study also has several limitations. First, this was a small sample of well-educated, predominantly White participants, which limits the generalizability of findings to other demographic populations. Second, there was no control arm in this pilot study meant to test feasibility, look for signals in clinical outcomes, and inform subsequent studies. This limitation prevents us from concluding that the observed stress reduction was due to the VR itself and not due to other confounding factors, such as the novelty of trying VR (for 11 participants). We plan to conduct future studies with a larger, more diverse population and incorporate a control group. Third, the design could not account for several factors such as time of day, caffeine intake, exercise, and medication use. This may significantly limit the measurement of physiological data and comparisons between participants [56]. While these patient-level factors were asked about in surveys (Multimedia Appendix 1),

we did not explicitly instruct participants to adjust their caffeine intake or exercise prior to the study visit, and we did not adjust for these factors in our analysis. Future studies should attempt to control for these considerations. Finally, this was also a 1-time intervention lacking a longer follow-up period to fully characterize long-term effects.

### Future Directions

Our findings warrant further research in a possible VR3 study, a randomized controlled trial that compares outcomes of an intervention and a control condition. Further, there are implications for designing a future study with a more specific population, more robust or targeted intervention, as well as for determining an appropriate control arm. There are several possibilities. One option would be to focus on a younger patient population at high clinical risk. Other options would be to focus on hospitalized patients during the periprocedural time period or in the intensive care unit. Alternatively, patients in CR may stand to benefit greatly due to depression and anxiety being highly prevalent in this population [80]. Stress reduction is one of the 3 main goals of CR. This VR intervention could be used in collaboration with a cardiac psychologist to target processing anxiety or trauma after a cardiac event, given that we found the VR experience activated certain memories and feelings in these individuals. Repeated use of the VR experience could also be considered. In terms of creating a control arm, possibilities include a traditional mindfulness meditation, progressive muscle relaxation, currently available VR programs for relaxation, an audio-only experience, or even allowing participants to try to take a nap. It will also be important to discuss possible VR improvements with designers such as using a custom headstrap mounting system to better support the weight of the display, shortening the length of the experience, changing the intensity of visuals, and potentially incorporating HRV biofeedback. Equally important will be continued partnership with patients from the specified populations to tailor VR content accordingly.

### Conclusions

This group of patients with or at risk for CVD exhibited higher-than-average stress levels, aligning with epidemiological findings. The notable reduction in perceived stress aligned with some but not all physiologic changes assessed. The VR intervention appeared to be a safe and practical method for stress reduction. Future studies are necessary to explore its effectiveness to lower stress in CVD at-risk and disease populations.

## Acknowledgments

The authors acknowledge Ms Dia Collins for her administrative support. The study would not have been possible without the collaboration and support from Ms Helena Choy and Dr Daniela Connolly.

## Data Availability

All data generated or analyzed during this study are included in this published paper (Multimedia Appendix 2) except for interview transcripts.

## Authors' Contributions

KEM designed the study, recruited participants, conducted study visits, analyzed all the data, and drafted and revised the manuscript. CV helped conduct study visits, analyzed data, and revised the manuscript. VG helped format data, conducted qualitative analysis, and revised the manuscript. LE recruited participants, conducted study visits, and revised the manuscript. MAC-P and KEW revised the manuscript. TH designed the study, oversaw study visits and analysis, and revised the manuscript.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Surveys, interview guide, longitudinal statistical testing results, and other survey results.
[DOCX File, 4262 KB - cardio_v9i1e66557_app1.docx ]

Multimedia Appendix 2
Physiologic and survey datasets generated and analyzed during this study.
[XLSX File, 170 KB - cardio_v9i1e66557_app2.xlsx ]

## References

1. Levine GN. The mind-heart-body connection. Circulation 2019 Oct 22;140(17):1363-1365. [doi: 10.1161/CIRCULATIONAHA.119.041914] [Medline: 31634016]
2. Levine GN, Cohen BE, Commodore-Mensah Y, et al. Psychological health, well-being, and the mind-heart-body connection: a scientific statement from the American Heart Association. Circulation 2021 Mar 9;143(10):e763-e783. [doi: 10.1161/CIR.0000000000000947]

3. Horwich T, Calfon-Press M, Watson KE. The heart-mind connection. Proc UCLA Health 2021;25 [FREE Full text]
4. Steptoe A, Kivimäki M. Stress and cardiovascular disease. Nat Rev Cardiol 2012 Apr 3;9(6):360-370. [doi: 10.1038/nrcardio.2012.45] [Medline: 22473079]
5. Vancheri F, Longo G, Vancheri E, Henein MY. Mental stress and cardiovascular health-Part I. J Clin Med 2022 Jun 10;11(12):3353. [doi: 10.3390/jcm11123353] [Medline: 35743423]
6. Song H, Fang F, Arnberg FK, et al. Stress related disorders and risk of cardiovascular disease: population based, sibling controlled cohort study. BMJ 2019;365:l1255. [doi: 10.1136/bmj.l1255]
7. Osborne MT, Shin LM, Mehta NN, Pitman RK, Fayad ZA, Tawakol A. Disentangling the links between psychosocial stress and cardiovascular disease. Circ: Cardiovascular Imaging 2020 Aug;13(8):e010931. [doi: 10.1161/CIRCIMAGING.120.010931]
8. Kubzansky LD, Huffman JC, Boehm JK, et al. Positive psychological well-being and cardiovascular disease. J Am Coll Cardiol 2018 Sep;72(12):1382-1396. [doi: 10.1016/j.jacc.2018.07.042]
9. Rugulies R. Depression as a predictor for coronary heart disease. a review and meta-analysis. Am J Prev Med 2002 Jul;23(1):51-61. [doi: 10.1016/s0749-3797(02)00439-7] [Medline: 12093424]
10. Yusuf S, Hawken S, Ôunpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004 Sep;364(9438):937-952. [doi: 10.1016/S0140-6736(04)17018-9]
11. Stewart RAH, Colquhoun DM, Marschner SL, et al. Persistent psychological distress and mortality in patients with stable coronary artery disease. Heart 2017 Dec;103(23):1860-1866. [doi: 10.1136/heartjnl-2016-311097] [Medline: 28652315]
12. Freedland KE, Carney RM. Depression as a risk factor for adverse outcomes in coronary heart disease. BMC Med 2013 May 15;11:131. [doi: 10.1186/1741-7015-11-131] [Medline: 23675637]
13. Mavrides N, Nemeroff CB. Treatment of affective disorders in cardiac disease. Dialogues Clin Neurosci 2015 Jun;17(2):127-140. [doi: 10.31887/DCNS.2015.17.2/nmavrides] [Medline: 26246788]
14. Rosengren A, Hawken S, Ounpuu S, et al. Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. Lancet 2004;364(9438):953-962. [doi: 10.1016/S0140-6736(04)17019-0] [Medline: 15364186]
15. Clemens B, Wagels L, Bauchmüller M, Bergs R, Habel U, Kohn N. Alerted default mode: functional connectivity changes in the aftermath of social stress. Sci Rep 2017 Jan 5;7:40180. [doi: 10.1038/srep40180] [Medline: 28054651]
16. Simon R, Engström M. The default mode network as a biomarker for monitoring the therapeutic effects of meditation. Front Psychol 2015;6:776. [doi: 10.3389/fpsyg.2015.00776] [Medline: 26106351]
17. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. Proc Natl Acad Sci U S A 2001 Jan 16;98(2):676-682. [doi: 10.1073/pnas.98.2.676] [Medline: 11209064]
18. Spiegel BMR. VRx: How Virtual Therapeutics Will Revolutionize Medicine: Basic Books; 2020.
19. Killingsworth MA, Gilbert DT. A wandering mind is an unhappy mind. Science 2010 Nov 12;330(6006):932. [doi: 10.1126/science.1192439] [Medline: 21071660]
20. Suzuki K, Roseboom W, Schwartzman DJ, Seth AK. A deep-dream virtual reality platform for studying altered perceptual phenomenology. Sci Rep 2017 Nov 22;7(1):15982. [doi: 10.1038/s41598-017-16316-2] [Medline: 29167538]
21. Rothbaum BO, Rizzo AS, Difede J. Virtual reality exposure therapy for combat-related posttraumatic stress disorder. Ann N Y Acad Sci 2010 Oct;1208:126-132. [doi: 10.1111/j.1749-6632.2010.05691.x] [Medline: 20955334]
22. Carl E, Stein AT, Levihn-Coon A, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J Anxiety Disord 2019 Jan;61:27-36. [doi: 10.1016/j.janxdis.2018.08.003]
23. Ong TL, Ruppert MM, Akbar M, et al. Improving the intensive care patient experience with virtual reality—a feasibility study. Crit Care Explor 2020 Jun;2(6):e0122. [doi: 10.1097/CCE.0000000000000122] [Medline: 32695991]
24. Failla C, Chilà P, Vetrano N, et al. Virtual reality for autism: unlocking learning and growth. Front Psychol 2024;15:1417717. [doi: 10.3389/fpsyg.2024.1417717] [Medline: 39691668]
25. Mistry D, Zhu J, Tremblay P, et al. Meditating in virtual reality: proof-of-concept intervention for posttraumatic stress. Psychol Trauma 2020 Nov;12(8):847-858. [doi: 10.1037/tra0000959] [Medline: 32897088]
26. Tinga AM, Nyklíček I, Jansen MP, de Back TT, Louwerse MM. Respiratory biofeedback does not facilitate lowering arousal in meditation through virtual reality. Appl Psychophysiol Biofeedback 2019 Mar;44(1):51-59. [doi: 10.1007/s10484-018-9421-5] [Medline: 30377895]
27. Chavez LJ, Kelleher K, Slesnick N, et al. Virtual reality meditation among youth experiencing homelessness: pilot randomized controlled trial of feasibility. JMIR Ment Health 2020 Sep 24;7(9):e18244. [doi: 10.2196/18244] [Medline: 32969834]
28. Highland KB, Costanzo ME, Jovanovic T, et al. Catecholamine responses to virtual combat: implications for post-traumatic stress and dimensions of functioning. Front Psychol 2015;6:256. [doi: 10.3389/fpsyg.2015.00256] [Medline: 25852586]
29. Roy MJ, Costanzo ME, Blair JR, Rizzo AA. Compelling evidence that exposure therapy for PTSD normalizes brain function. Stud Health Technol Inform 2014;199:61-65. [Medline: 24875691]
30. Birckhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973. [doi: 10.2196/11973] [Medline: 30702436]

31. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
32. Baik SH, Fox RS, Mills SD, et al. Reliability and validity of the Perceived Stress Scale-10 in Hispanic Americans with English or Spanish language preference. J Health Psychol 2019 Apr;24(5):628-639. [doi: 10.1177/1359105316684938] [Medline: 28810432]
33. Cohen S, Williamson G. Perceived stress in a probability sample of the U.S. In: Spacapam S, Oskamp S, editors. The Social Psychology of Health: Claremont Symposium on Applied Social Psychology: Sage Publications Inc; 1988.
34. Cohen S, Janicki‐deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J Applied Social Pyschol 2012 Jun;42(6):1320-1334. [doi: 10.1111/j.1559-1816.2012.00900.x]
35. Metzger RL. A reliability and validity study of the state-trait anxiety inventory. J Clin Psychol 1976 Apr;32(2):276-278. [doi: 10.1002/1097-4679(197604)32:2<276::AID-JCLP2270320215>3.0.CO;2-G]
36. Turrado V, Guzmán Y, Jiménez-Lillo J, et al. Exposure to virtual reality as a tool to reduce peri-operative anxiety in patients undergoing colorectal cancer surgery: a single-center prospective randomized clinical trial. Surg Endosc 2021 Jul;35(7):4042-4047. [doi: 10.1007/s00464-021-08407-z] [Medline: 33683433]
37. Kütmeç Yilmaz C, Duru Aşiret G, Çetinkaya F. The effect of back massage on physiological parameters, dyspnoea, and anxiety in patients with chronic obstructive pulmonary disease in the intensive care unit: a randomised clinical trial. Intensive Crit Care Nurs 2021 Apr;63:102962. [doi: 10.1016/j.iccn.2020.102962] [Medline: 33162314]
38. Kim J, Kim H, Kim DH, et al. Effects of cranial electrotherapy stimulation with novel in-ear electrodes on anxiety and resting-state brain activity: a randomized double-blind placebo-controlled trial. J Affect Disord 2021 Dec 1;295:856-864. [doi: 10.1016/j.jad.2021.08.141] [Medline: 34706456]
39. Julian LJ. Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11(11):S467-S472. [doi: 10.1002/acr.20561] [Medline: 22588767]
40. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory: Consulting Psychologists Press; 1983.
41. Gilgen-Ammann R, Schweizer T, Wyss T. RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. Eur J Appl Physiol 2019 Jul;119(7):1525-1532. [doi: 10.1007/s00421-019-04142-5] [Medline: 31004219]
42. Speer KE, Semple S, Naumovski N, McKune AJ. Measuring heart rate variability using commercially available devices in healthy children: a validity and reliability study. Eur J Investig Health Psychol Educ 2020 Jan 10;10(1):390-404. [doi: 10.3390/ejihpe10010029] [Medline: 34542492]
43. Müller AM, Wang NX, Yao J, et al. Heart rate measures from wrist-worn activity trackers in a laboratory and free-living setting: validation study. JMIR Mhealth Uhealth 2019 Oct 2;7(10):e14120. [doi: 10.2196/14120] [Medline: 31579026]
44. Weaver L, Wooden T, Grazer J. Validity of Apple watch heart rate sensor compared to Polar H10 heart rate monitor [Georgia College and State University]. J Stud Res 2019. [doi: 10.47611/jsr.vi.662]
45. Hinde K, White G, Armstrong N. Wearable devices suitable for monitoring twenty four hour heart rate variability in military populations. Sensors (Basel) 2021 Feb 4;21(4):1061. [doi: 10.3390/s21041061] [Medline: 33557190]
46. Perrotta AS, Jeklin AT, Hives BA, Meanwell LE, Warburton DER. Validity of the Elite HRV smartphone application for examining heart rate variability in a field-based setting. J Strength Cond Res 2017 Aug;31(8):2296-2302. [doi: 10.1519/JSC.0000000000001841] [Medline: 28195974]
47. Telles S, Nagarathna R, Nagendra HR. Autonomic changes while mentally repeating two syllables—one meaningful and the other neutral. Indian J Physiol Pharmacol 1998 Jan;42(1):57-63. [Medline: 9513794]
48. Singh Y, Sharma R, Talwar A. Immediate and long-term effects of meditation on acute stress reactivity, cognitive functions, and intelligence. Altern Ther Health Med 2012;18(6):46-53. [Medline: 23251943]
49. Bharshankar JR, Mandape AD, Phatak MS, Bharshankar RN. Autonomic functions in Raja-yoga meditators. Indian J Physiol Pharmacol 2015;59(4):396-401. [Medline: 27530006]
50. Flagler T, Tong J, Allison RS, Wilcox LM. Validity testing the neulog galvanic skin response device. Presented at: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC); Oct 11-14, 2020; Toronto, ON, Canada p. 3964-3968. [doi: 10.1109/SMC42975.2020.9283465]
51. Witmer BG, Singer MJ. Measuring presence in virtual environments: a presence questionnaire. Presence (Camb) 1998 Jun;7(3):225-240. [doi: 10.1162/105474698565686]
52. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
53. Bimberg P, Weissker T, Kulik A. On the usage of the simulator sickness questionnaire for virtual reality research. Presented at: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); Mar 22-26, 2020; Atlanta, GA, USA p. 464-467. [doi: 10.1109/VRW50115.2020.00098]
54. Lipponen JA, Tarvainen MP. A robust algorithm for heart rate variability time series artefact correction using novel beat classification. J Med Eng Technol 2019 Apr;43(3):173-181. [doi: 10.1080/03091902.2019.1640306] [Medline: 31314618]

55. Stone JD, Ulman HK, Tran K, et al. Assessing the accuracy of popular commercial technologies that measure resting heart rate and heart rate variability. Front Sports Act Living 2021;3:585870. [doi: 10.3389/fspor.2021.585870] [Medline: 33733234]
56. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health 2017;5:258. [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]
57. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation 1996;93. [Medline: 859806]
58. Sloan J, Symonds T, Vargas-Chanes D, Fridley B. Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. Drug Information J 2003 Jan;37(1):23-31. [doi: 10.1177/009286150303700105]
59. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003 May;41(5):582-592. [doi: 10.1097/01.MLR.0000062554.74615.4C] [Medline: 12719681]
60. Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. J Clin Epidemiol 2005 Dec;58(12):1217-1219. [doi: 10.1016/j.jclinepi.2005.07.009]
61. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological: American Psychological Association:57-71.
62. Creswell JW, Poth CN. Qualitative Inquiry and Research Design Choosing among Five Approaches, 4th edition: SAGE Publications Inc; 2018.
63. Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed method designs in implementation research. Adm Policy Ment Health 2011 Jan;38(1):44-53. [doi: 10.1007/s10488-010-0314-z] [Medline: 20967495]
64. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience: Harper & Row; 1990.
65. Kotler S, Wheal J. Stealing Fire: How Silicon Valley, the Navy SEALS, and Maverick Scientists Are Revolutionizing the Way We Live and Work: HarperCollins; 2017:28-36.
66. Dietrich A. Functional neuroanatomy of altered states of consciousness: the transient hypofrontality hypothesis. Conscious Cogn 2003 Jun;12(2):231-256. [doi: 10.1016/s1053-8100(02)00046-6] [Medline: 12763007]
67. Speed C, Arneil T, Harle R, et al. Measure by measure: resting heart rate across the 24-hour cycle. PLOS Digit Health 2023 Apr;2(4):e0000236. [doi: 10.1371/journal.pdig.0000236] [Medline: 37115739]
68. Li K, Rüdiger H, Ziemssen T. Spectral analysis of heart rate variability: time window matters. Front Neurol 2019;10:545. [doi: 10.3389/fneur.2019.00545] [Medline: 31191437]
69. Garg P, Mendiratta A, Banga A, et al. Effect of breathing exercises on blood pressure and heart rate: a systematic review and meta-analysis. Int J Cardiol Cardiovasc Risk Prev 2024 Mar;20:200232. [doi: 10.1016/j.ijcrp.2023.200232] [Medline: 38179185]
70. Najafpour E, Asl-Aminabadi N, Nuroloyuni S, Jamali Z, Shirazi S. Can galvanic skin conductance be used as an objective indicator of children's anxiety in the dental setting? J Clin Exp Dent 2017 Mar;9(3):e377-e383. [doi: 10.4317/jced.53419] [Medline: 28298978]
71. Garrido LE, Frías-Hiciano M, Moreno-Jiménez M, et al. Focusing on cybersickness: pervasiveness, latent trajectories, susceptibility, and effects on the virtual reality experience. Virtual Real 2022;26(4):1347-1371. [doi: 10.1007/s10055-022-00636-4] [Medline: 35250349]
72. Dużmańska N, Strojny P, Strojny A. Can simulator sickness be avoided? A review on temporal aspects of simulator sickness. Front Psychol 2018;9:2132. [doi: 10.3389/fpsyg.2018.02132] [Medline: 30459688]
73. Kim H, Kim DJ, Kim S, et al. Effect of virtual reality on stress reduction and change of physiological parameters including heart rate variability in people with high stress: an open randomized crossover trial. Front Psychiatry 2021;12:614539. [doi: 10.3389/fpsyt.2021.614539] [Medline: 34447320]
74. Baytar AD, Bollucuo Lu K. Effect of virtual reality on preoperative anxiety in patients undergoing septorhinoplasty. Braz J Anesthesiol 2023;73(2):159-164. [doi: 10.1016/j.bjane.2021.08.014] [Medline: 34562488]
75. Niki K, Yahara M, Inagaki M, et al. Immersive virtual reality reminiscence reduces anxiety in the oldest-old without causing serious side effects: a single-center, pilot, and randomized crossover study. Front Hum Neurosci 2020;14:598161. [doi: 10.3389/fnhum.2020.598161] [Medline: 33536887]
76. Brown K, Foronda C. Use of virtual reality to reduce anxiety and pain of adults undergoing outpatient procedures. Informatics (MDPI) 2020;7(3):36. [doi: 10.3390/informatics7030036]
77. Karaman D, Taşdemir N. The effect of using virtual reality during breast biopsy on pain and anxiety: a randomized controlled trial. J Perianesth Nurs 2021 Dec;36(6):702-705. [doi: 10.1016/j.jopan.2021.04.007]
78. Morgan H, Nana M, Phillips D, Gallagher S. The Effect of a VIRTUAL RealiTy Immersive Experience Upon Anxiety Levels, Procedural Understanding, and Satisfaction in Patients Undergoing CArdiac CaTHeterization: The VIRTUAL CATH Trial. J Invasive Cardiol 2021 Sep;33(9):E681-E686. [doi: 10.25270/jic/20.00664] [Medline: 34398809]
79. Keshvari M, Yeganeh MR, Paryad E, Roushan ZA, Pouralizadeh M. The effect of virtual reality distraction on reducing patients' anxiety before coronary angiography: a randomized clinical trial study. Egypt Heart J 2021 Nov 4;73(1):98. [doi: 10.1186/s43044-021-00224-y] [Medline: 34735643]

80. Bermudez T, Bierbauer W, Scholz U, Hermann M. Depression and anxiety in cardiac rehabilitation: differential associations with changes in exercise capacity and quality of life. Anxiety Stress Coping 2022 Mar;35(2):204-218. [doi: 10.1080/10615806.2021.1952191] [Medline: 34269151]

## Abbreviations

**BP:** blood pressure
**CR:** cardiac rehabilitation
**CVD:** cardiovascular disease
**DMN:** default mode network
**GSR:** galvanic skin response
**HF:** high frequency
**HR:** heart rate
**HRV:** heart rate variability
**ITQ:** Immersive Tendencies Questionnaire
**LF:** low frequency
**PSS-10:** Perceived Stress Scale-10
**RMSSD:** root-mean-square of successive differences
**SDNN:** standard deviation of the interbeat interval of normal sinus beats
**SSQ:** Simulation Sickness Questionnaire
**STAI-S:** State-Trait Anxiety Inventory-State
**UCLA:** University of California Los Angeles
**VR:** virtual reality

*Edited by A Coristine; submitted 16.09.24; peer-reviewed by S Mitra, T Cahill; revised version received 17.03.25; accepted 17.03.25; published 06.08.25.*

*Please cite as:*
*Makaroff KE, Van C, Grospe V, Edmunds L, Calfon-Press MA, Watson KE, Horwich T*
*Novel Virtual Reality Intervention for Stress Reduction Among Patients With or at Risk for Cardiovascular Disease: Mixed Methods Pilot Study*
*JMIR Cardio 2025;9:e66557*
*URL: https://cardio.jmir.org/2025/1/e66557*
*doi:10.2196/66557*

# Long-Term Feasibility and Outcomes of a Digital Health Program to Improve Liver Fat and Cardiometabolic Markers in Individuals With Nonalcoholic Fatty Liver Disease: Prospective Single-Arm Feasibility Study

Sigridur Björnsdottir[1,2], MD, PhD; Hildigunnur Ulfsdottir[3], MD, MPH; Elias Freyr Gudmundsson[3], BSc, MSc; Bartosz Dobies[3,4], MSc; Kolbrun Sveinsdottir[3], MSc; Ari Pall Isberg[3], MSc; Gudlaug E A Magnusdottir[3], BSc; Thrudur Gunnarsdottir[3], PhD; Tekla Karlsdottir[3], MD; Gudlaug Bjornsdottir[5,6], MSc; Sigurdur Sigurdsson[5,6], PhD; Saemundur Oddsson[3], MD; Vilmundur Gudnason[5,6], MD, PhD

[1]Department of Molecular Medicine and Surgery, Karolinska University Hospital, Eugeniavägen 3, Solna, Stockholm, Sweden

[2]Heilsuklasinn, Reykjavik, Iceland

[3]Clinical Office, Sidekick Health, Vallakor 4, Kopavogur, Iceland

[4]Doctoral School of Medical and Health Sciences, Faculty of Medicine, Jagiellonian University Medical College, Krakow, Poland

[5]Icelandic Heart Association, Kopavogur, Iceland

[6]School of Health Sciences, Faculty of Medicine, University of Iceland, Reykjavik, Iceland

**Corresponding Author:**
Sigridur Björnsdottir, MD, PhD
Department of Molecular Medicine and Surgery, Karolinska University Hospital, Eugeniavägen 3, Solna, Stockholm, Sweden

## Abstract

**Background:** A 12-week digital health program for nonalcoholic fatty liver disease (NAFLD) previously showed feasibility in engagement, program retention, and clinical outcomes. This study investigates whether improvements in cardiometabolic risk factors achieved during a 12-week active program were sustained over a subsequent 6-month follow-up period.

**Objective:** The primary objective of this analysis was to evaluate whether the clinical improvements achieved after a 12-week program were maintained over the subsequent 6-month period, which did not include coaching or new intervention materials. In addition, the study aimed to assess participants' retention and engagement with the maintenance program.

**Methods:** In a 9-month, single-arm study using the Sidekick app (Sidekick Health), individuals with NAFLD and BMI >30 or metabolic syndrome or type 2 diabetes were included. The initial 12 weeks focused on providing education about diet, physical activity, stress management, and sleep, followed by 6 months without coaching or new intervention materials. The measured outcomes encompassed demographics, body composition, liver fat assessed using magnetic resonance imaging-proton density fat fraction (MRI-PDFF), and blood markers.

**Results:** Of the 34 participants who completed the first 12 weeks, 28 (82%) completed the 9-month study measurements. The median age was 63.0 years (IQR 53.5 - 71.0) and 57.1% (16/28) were women. At 9 months, compared to baseline, the mean weight loss was 4.0 kg (SD 5.0; *P*<.001). Liver fat decreased by 2.5% (SD 4.5; *P*<.001), with an 18.4% relative reduction. Systolic blood pressure decreased by 8.3 mm Hg (SD 13.4, *P*<.001) and diastolic by 2.5 mm Hg (SD 6.0; *P*=.02). Waist circumference decreased by 4.7 cm (SD 7.1; *P*<.001) and median glycated hemoglobin A1c ($HbA_{1c}$) decreased by 19.5 mmol/mol (*P*<.001).

**Conclusion:** Sustained improvements in liver fat and metabolic markers suggest that Sidekick Health's digital program is a promising strategy for managing NAFLD without requiring continuous coaching.

**Trial Registration:** ClinicalTrials.gov NCT05426382; https://clinicaltrials.gov/study/NCT05426382

***(JMIR Cardio 2025;9:e72074)*** doi:10.2196/72074

**KEYWORDS**

digital health program; nonalcoholic fatty liver disease; NAFLD; cardiometabolic health; digital therapeutics; liver; chronic; hepatic; cardiometabolic; cardiovascular; cardiology; weight; acceptability; digital health; metabolic syndrome; diabetic; diabetes; type 2 diabetes; BMI; lifestyle; exercise; physical activity; coaching; diet; dietary; nutrition; nutritional; patient education; coach; feasibility; fat; body composition; MAFLD; MASLD

## Introduction

As lifestyles have become more sedentary and diets have become hypercaloric, with high levels of refined sugars, grains, ultraprocessed foods, and sugar-containing beverages, nonalcoholic fatty liver disease (NAFLD) has emerged as a major public health concern. This rise parallels the global increase in obesity, type 2 diabetes, and metabolic syndrome [1,2]. NAFLD is now the most prevalent chronic liver disease globally, characterized by the accumulation of fat in the liver in the absence of excessive alcohol consumption or other identifiable secondary causes [3]. NAFLD encompasses a spectrum of liver conditions, from simple steatosis to its more severe form, nonalcoholic steatohepatitis (NASH), which can progress to fibrosis, cirrhosis, liver failure, and hepatocellular carcinoma [4]. NAFLD is a multisystem disease that is closely interlinked with metabolic syndrome, and they share common pathophysiological mechanisms and frequently coexist. NAFLD not only represents the hepatic manifestation of metabolic syndrome but may also contribute to its progression, establishing a mutual and bidirectional relationship between the 2 conditions [5,6]. In individuals with NAFLD, the leading cause of mortality is cardiovascular disease, followed by mortality from extrahepatic cancer, liver-related conditions (including hepatocellular carcinoma), and diabetes [7]. Furthermore, it is predicted that NAFLD and NASH will soon become the most common indication for liver transplantation, highlighting the growing socioeconomic burden associated with this disease [8].

Despite the rising prevalence of NAFLD, effective therapeutic options are still limited, with no approved pharmacological treatment currently available for its management [9]. Currently, lifestyle modifications focusing on dietary changes, increased physical activity, weight loss, and weight maintenance are the primary initial approaches for managing NAFLD [10]. However, the lack of standardized, evidence-based interventions and limited out-of-hospital monitoring often leads to suboptimal long-term results for individuals undergoing lifestyle interventions for NAFLD. This emphasizes the urgent need for innovative and scalable solutions to mitigate disease progression and improve long-term clinical outcomes. The use of digital technologies, such as smartphone apps, offers a promising approach to providing scalable and personalized interventions for individuals with NAFLD, with encouraging results reported in a recent study [11].

Sidekick Health is an Icelandic digital therapeutic company that has developed a digital health program (Sidekick-241 or SK-241) specifically designed for people with NAFLD and metabolic derangements. The emphasis of the SK-241 program is to improve nutritional status by focusing on limiting ultraprocessed foods, decreasing carbohydrates, increasing physical activity levels, reducing stress, and improving sleep. The results from the initial 12-week active phase of the SK-241 digital health program have been previously published, showing excellent retention, engagement, and satisfaction, alongside improvements in liver-specific and cardiometabolic health outcomes [12]. In this study, we present the 9-month results, including the 6-month maintenance period, highlighting the potential longer-term benefit of the SK-241 digital health program.

## Methods

### Trial Design

This open-label, single-arm, prospective study was conducted over 9 months, from June 2022 to April 2023 in Iceland. The study period included an active 12-week digital health program delivered through the Sidekick app, followed by an optional 6-month maintenance period. Clinical assessments, including screening and pre- and postprogram clinical assessments, were carried out at baseline, 12 weeks, and 9 months at The Icelandic Heart Association.

### Nomenclature

Recent updates in the nomenclature for fatty liver disease have introduced the term Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD) as a replacement for NAFLD [13]. While this new terminology has been endorsed by several major liver societies, it has not yet reached universal acceptance and has sparked some debate and confusion within the scientific community [14,15]. Given that the diagnostic criteria and study design for the present study were developed and registered under the established NAFLD framework, and to maintain consistency with previous literature, we have retained the use of the term NAFLD in this manuscript. This decision aligns with current guidance suggesting that researchers may continue to use the nomenclature most appropriate to their study context during this transitional period [16]. Future studies will adopt the updated consensus terminology as standardization across the field progresses.

### Participants and Screening for NAFLD

The screening process has previously been described [12]. Briefly, 38 individuals between 18 - 80 years old, with at least 1 of the following risk factors: BMI >30, metabolic syndrome or type 2 diabetes mellitus, and confirmed liver steatosis >5% with a noninvasive ultrasonography-based controlled attenuation parameter (CAP) assessment through a FibroScan device were invited to participate after giving informed consent and fulfilling the inclusion and exclusion criteria [17].

### Digital Health Program

The active 12-week digital health program has previously been described [12]. To summarize, the primary focus of the 12-week-long program was to reduce participants' daily dietary carbohydrate consumption and improve their overall nutrition quality in small, achievable, and sustainable steps (eg, reducing added sugars and processed foods, prioritizing protein, and increasing vegetable consumption). A secondary focus was to increase daily physical activity levels, improve sleep quality, and reduce stress. The program included short daily missions aimed at increasing knowledge about NAFLD and its contributing factors. The daily missions included watching short educational videos, reading brief informational content, and logging meals and beverages, which consisted of taking a photo of the meal, assessing on a sliding scale how healthy the meal was, and evaluating hunger and satiety before and after the meal.

Other missions were regular practice of mindfulness and meditation, logging daily energy levels, stress, and sleep quality on a sliding scale from 0 - 10. The 12-week active program also provided participants with in-app health coach support (a person, not artificial intelligence [AI]), who gave weekly feedback on food logs and other in-app activities and answered participants' questions as needed. In addition, at the beginning of the 12-week program, participants had the opportunity to have a 30-minute video call interview with the health coach for a baseline assessment.

After completing the 12-week program, participants were given the option to retain access to the app for an additional 6 months. During this maintenance period, no new content was provided; however, participants could access previously received educational materials and continue using features such as a food-logging tool, in-app step counter, and other features. Active health coach support was not included during the 6-month maintenance period.

## Outcome Measures and Covariates

The primary objective of this analysis was to evaluate whether the clinical improvements achieved after a 12-week program were maintained over the subsequent 6-month period, which did not include coaching or new intervention materials. In addition, the study aimed to assess participants' retention and engagement with the maintenance program.

Participants were assessed at baseline, 12-week, and 9-month for demographic data, anthropometric measures, medical history, medications, and adverse events. Liver fat content was measured and quantified at these time points using MRI-PDFF with a multiecho chemical shift-encoded gradient-echo sequence [18]. A new MRI machine was used for the 9-month follow-up visit. Analysis confirmed that the 2 machines produced comparable results, with an intraclass correlation coefficient of 0.97 (Multimedia Appendix 1). Body composition was assessed at baseline, 12 weeks, and 9 months, with a dual-energy x-ray absorptiometry (DXA) [19]. Blood pressure was measured using an automatic blood pressure monitor. Blood samples were drawn at the same time points to measure complete blood count, alanine aminotransferase (ALAT), aspartate aminotransferase (ASAT), glycated hemoglobin A1c ($HbA_{1c}$), fasting glucose and insulin for Homeostatic Model Assessment of Insulin Resistance (HOMA-IR), total cholesterol, high-density lipoprotein-cholesterol (HDL-C), low-density lipoprotein-cholesterol (LDL-C), triglycerides, and high-sensitivity C-reactive protein (hs-CRP). The Fibrosis-4 Index, a noninvasive tool to estimate liver fibrosis risk and categorize outcomes in low-, medium-, and high-risk fibrosis groups, was calculated at baseline, 12 weeks, and 9 months [20,21].

Participants were given the following questionnaires via an electronic Patient Reported Outcome (PRO): Depression, Anxiety, and Stress Scale (DASS-21), health-related quality of life (EQ-5D-5L), and the Morisky Medication Adherence Scale (MMAS-8) [22-25].

For an exploratory engagement analysis, study participants were divided into 2 groups depending on how engaged they were with the digital health program to assess if clinical outcomes were associated with in-app use and engagement. For the 6-month maintenance period, clinical outcomes were compared between those participants who remained active during the maintenance period (defined as being active 18 or more weeks of the 24 weeks) and those who were less active.

For a supplementary exploratory behavior change analysis, participants were administered an in-app questionnaire at baseline and again at week 12, focusing on behavior change. To analyze the pre-post responses, the answers were categorized into 2 groups (Yes or No, Disagree or Agree, Rarely True, Often True) and the frequency of each category was reported. The analysis included those who had answered the questionnaire at both baseline and week 12. The McNemar chi-square test with continuity correction was applied to evaluate the significance of changes in response rates from baseline to week 12.

## Statistical Analysis

A formal sample size calculation was not performed as this was a feasibility study. It was considered sufficient to aim for 30 - 40 participants to obtain information on practical aspects of recruitment, in-app engagement, retention, and rates of acceptance, which were the primary outcomes of the study.

Changes in clinical assessments and PROs between follow-up time points were calculated as the mean and SD for approximately normally distributed variables (normality was analyzed with the Shapiro-Wilk test) or as the median and IQR for variables that did not satisfy normality criteria. Categorical data were calculated as frequencies and percentages. To compare baseline and follow-up outcomes, paired *t* tests were computed for approximately normally distributed data. In case the normality assumption was not met, nonparametric tests were computed (Wilcoxon signed-rank tests). Unless otherwise specified, all statistical tests were performed at the 5% (2-sided) significance level. Statistical analysis was performed in Stata (StataCorp) and R (version 4.0.3; R Foundation for Statistical Computing). All enrolled participants were included in the full analysis set. Missing data were imputed using the baseline observation carried forward provided that the participant was enrolled in the study and at least 1 of 3 measurements (baseline or follow-up at 12 weeks or 9 months) was collected. Moreover, missing baseline measurements in waist circumference, hip circumference, and low-density lipoprotein cholesterol were imputed for 1 participant using the next observation carried backward. The complete case analysis set included participants who attended both the baseline visit and the 12-week follow-up visit.

## Ethical Considerations

The National Bioethics Committee of Iceland and the Data Protection Authority approved this study under the approval code 22 - 075-Vl. All participants provided informed consent before enrolling in the study. All data were deidentified and analyzed in accordance with institutional protocols. Participants were given the option of seeking reimbursement for travel expenses not exceeding US $150 in total; no other compensation was provided. The study was registered at clinicaltrials.gov under the trial identifier NCT05426382.

## Results

### Participant Characteristics

A total of 28 participants completed the measurements following the maintenance program. The median age of those who completed was 63.0 years (IQR 53.5 - 71.0), with 16 (57%) being female, and all participants identifying as Caucasian. None of the participants were smokers, and 42% held a university degree. At baseline, 89% of participants had obesity (BMI >30), 60% had type 2 diabetes mellitus, 75% had hypertension, 46% had hypercholesterolemia, and 40% had a history of cardiovascular disease. In total, 53% of participants reported taking antidiabetic medication, 85% antihypertensive medication, and 46% antilipidemic medication. During the 6-month maintenance period, 16 participants (42%) reported changes to their medication: 16 (42%) started new medications, 4 (10.5%) had dosage adjustments (in strength and or frequency), and 7 (18%) reported discontinuing medication. Of particular interest were changes that could influence metabolic outcomes. One participant started a glucagon-like peptide-1 agonist (GLP1-RA), semaglutide, along with metformin. Another participant switched from semaglutide to liraglutide, 1 increased their dose of semaglutide, and 1 participant stopped taking semaglutide. Other medication changes were considered irrelevant to the study outcomes by the study's principal investigator.

### Retention and Engagement in the Maintenance Period

Of the 34 individuals who completed the 12-week active program, 28 (82%) individuals attended the third and final follow-up visit at 9 months. By the end of the 6-month maintenance period, 19 out of 38 participants (50%) were still retained in the app during the final week. In addition, 17 participants (45%) were active >5 days during 18 of the 24 weeks of the maintenance period. The median number of active days per week during this period was 2.38 (IQR 0.36 - 6.08) days, and participants completed an average of 4.2 (SD 5.9) daily missions with the app per day.

### Metabolic Parameters

At month 9 compared to baseline (Table 1), participants in the full analysis set (n=38) demonstrated significant improvement in metabolic parameters.

The mean weight loss (SD) was 4.0 (5.0) kg ($P$<.001). The mean (SD) absolute reduction in liver fat was 2.5% (4.5) and the mean (SD) relative reduction was 18.4% (30.5) ($P$<.001). There was also a reduction in mean (SD) systolic blood pressure by 8.3 mm Hg (13.4) ($P$<.001) and in diastolic blood pressure by 2.5 mm Hg (6.0) ($P$=.02). The mean (SD) waist circumference decreased by 4.7 cm (7.1) ($P$<.001) and the median (IQR) $HbA_{1c}$ was reduced by 18.5 mmol/mol (3-22) ($P$<.001).

When comparing the first 12 weeks to the subsequent 6 months maintenance period, all metabolic parameters remained significantly improved (Table 1), except for triglycerides ($P$=.15) and a slight increase in LDL-cholesterol. The mean (SD) LDL-cholesterol value was 2.9 mmol/L (1.1) at week 12 compared to 3.0 mmol/L (1.0) at month 9 ($P$=.04). In addition, diastolic blood pressure continued to decline during the maintenance period, resulting in a significant mean (SD) reduction of 2.5 mm Hg (6.0; $P$=.02) at month 9, which was not observed at week 12. These improvements were not explained by changes in medication or medication adherence (data not shown). Furthermore, the median (IQR) hs-CRP value significantly decreased from 3.0 mg/L (1.2 - 5.2) at baseline to 2.4 mg/L (1.1 - 3.9; $P$=.03) at month 9. In contrast, the reduction in hs-CRP at week 12 was not statistically significant. At baseline, the mean fasting s-insulin and HOMA-IR levels indicated insulin resistance. At week 12, we detected significant improvements in these variables and found that those improvements were sustained over the maintenance period (see Table 1). In addition, we saw further improvements in glycemic control during the maintenance period, with a median (IQR) reduction of 18.5 mmol/mol (3-22) ($P$<.001) in $HbA_{1c}$ levels, compared with week 12. These data indicate that initial clinical improvements were sustained over the maintenance period.

There was a trend to an even better improvement in most of the metabolic parameters at 9 months compared to 12 weeks, although this was not significant.

The sustained improvements in weight loss and body composition were mirrored by sustained reduction in liver fat measured by MRI-PDFF. The mean (SD) liver fat percentage was 9.8% (6.6) at month 9 and 10.1% (6.5) at week 12. The mean (SD) relative reduction from baseline in MRI-PDFF liver fat value at month 9 was 18.4% (30.5), aligning closely with the reduction observed during the first 12 weeks of the study ($P$<.001).

**Table .** Differences in anthropometric, biochemical, and clinical measurements at baseline, week 12, and month 9 for the full analysis set.

| Characteristics | Baseline (n=38) | 12-week (n=38) | Month 9 (n=38) | Change from baseline to Week 12 | Change from baseline to Month 9 | Change from Week 12 to Month 9 | *P* value baseline versus Week 12[a] | *P* value baseline versus Month 9[a] | *P value* Week 12 versus Month 9[a] |
|---|---|---|---|---|---|---|---|---|---|
| Anthropometry | | | | | | | | | |
| Weight (kg), mean (SD) | 110.0 (18.5) | 106.5 (18.4) | 106.0 (19.4) | 3.5 (3.7) | 4.0 (5.0) | 0.5 (4.4) | <.001 | <.001 | .48 |
| Relative % change in weight, mean (SD) | N/A[b] | N/A[b] | N/A[b] | 3.2 (3.4) | 3.8 (4.9) | 0.5 (0.4) | N/A[b] | N/A[b] | N/A[b] |
| BMI (kg/m2), mean (SD) | 37.6 (5.8) | 36.4 (5.8) | 36.2 (6.1) | 1.2 (1.3) | 1.3 (1.7) | 0.2 (1.4) | <.001 | <.001 | .46 |
| Waist circumference (cm), mean (SD) | 123.8 (12.2) | 119.9 (12.2) | 118.9 (13.7) | 4.0 (5.1) | 4.7 (7.1) | 0.7 (5.5) | <.001 | <.001 | .45 |
| Hip circumference (cm), mean (SD) | 125.1 (14.0) | 123.2 (13.3) | 122.6 (14) | 1.8 [0.0 - 4.9] | 2.5 (6.7) | 0.6 (5.0) | .01 | .03 | .45 |
| Waist to hip ratio, median, (IQR) | 1.00 [0.95 - 1.03] | 0.99 [0.92 - 1.03] | 1 [0.9 - 1.0] | 0.00 [−0.01 - 0.03] | 0.02 (0.06) | 0.0 [0.0 - 0.0] | .09[d] | .02[d] | .33[d] |
| Liver assessment | | | | | | | | | |
| Liver fat MRI-PDFF (%), mean (SD) | 12.3 (7.1) | 10.1 (6.5) | 9.8 (6.6) | 2.2 (2.9) | 2.5 (4.5) | 0.3 (3.9) | <.001 | <.001 | .60 |
| Liver fat MRI-PDFF mean relative change (%)[c] | N/A[b] | N/A[b] | | 19.4 (23.9) | 18.4 (30.5) | 5.8 (0.47) | N/A[b] | N/A[b] | N/A[b] |
| Liver stiffness measure (kPa), median (IQR) | 6.4 [5.2 - 9.6] | 6.6 [5.3 - 8.4] | 6.4 [5.0 - 8.4] | 0.2 [−0.3 - 1.6] | 0.0 [−0.1 - 1.2] | 0.2 [−0.4 - 1.8] | .11[d] | .11[d] | .15[d] |
| CAP score (dB/m), mean (SD)[e] | 343.6 (34.8) | 310.3 (47.2) | 315.2 (54.3) | 33.3 (39.7) | 16.0 [0.0 - 50.0] | −4.9 (51.6) | <.001 | <.001 | .56 |
| Body composition | | | | | | | | | |
| Total body region fat, %, median (IQR)[f] | 46.6 [39.4 - 52.4] | 44.3 [37.8 - 52.2] | 44.4 [8.1 - 51.4] | 0.9 (1.4)[c] | 1.0 (2.6) | 0.1 (2.4) | <.001 | <.001 | .81 |
| Fat mass, kg [f] | 50.3 (13.8) | 48.1 (14.5) | 47.8 (14.5) | 2.2 (2.7) | 2.4 (3.9) | 0.2 (3.7) | <.001 | <.001 | .72 |
| Lean mass, kg[f] | 56.3 (10.1) | 55.6 (9.7) | 55.3 (8.9) | 0.7 (1.7) | 1.0 (2.4) | 0.3 (2.3) | .008 | <.001 | .42 |

| Characteristics | Baseline (n=38) | 12-week (n=38) | Month 9 (n=38) | Change from baseline to Week 12 | Change from baseline to Month 9 | Change from Week 12 to Month 9 | *P* value baseline versus Week 12[a] | *P* value baseline versus Month 9[a] | *P value* Week 12 versus Month 9[a] |
|---|---|---|---|---|---|---|---|---|---|
| Blood pressure (mm Hg), mean (SD) | | | | | | | | | |
| Systolic | 141.4 (17.1) | 135.4 (17.3) | 133.1 (14.7) | 6.0 (13.5) | 8.3 (13.4) | 2.3 (13.0) | .009 | <.001 | .29 |
| Diastolic | 83.6 (7.4) | 82.5 (7.4) | 81.1 (7.9) | 1.2 (7.7) | 2.5 (6.0) | 1.3 (8.1) | .36 | .02 | .32 |
| Biochemical measures | | | | | | | | | |
| $HbA_{1c}$ (mmol/mol), median [IQR] [g] | 60.0 [56.0 - 66.8] | 60.0 [54.3 - 64.0] | 42.5 (39-55) | 0.5 [−0.7 - 3.8] | 19.5 [2-22] | 18.5 [3-22] | .03[d] | <.001[d] | **<.001**[d] |
| S-Glucose (mmol/L), median [IQR][h] | 6.2 [5.3 - 7.4] | 6.3 [5.4 - 6.9] | 6.3 [5.6 - 6.8] | 0.0 [−0.3 - 0.4] | 0 [−.2 - 0.1] | 0.0 [−0.4-.4.0] | .64[d] | .76[d] | .94[d] |
| S-Insulin (µU/ml), median [IQR][i] | 21.1 [16.4 - 27.9] | 19.0 [13.0 - 25.0] | 18.9 [13.5 - 24.5] | 3.2 [0.0 - 5.4] | 1.7 [0 - 6.5] | 0.0 [-3.4 - 3.1] | .003[d] | <.001[d] | .64[d] |
| HOMA-IR (mmol/L), median (IQR)[j] | 5.8 [4.3 - 8.4] | 4.8 [3.6 - 7.2] | 5.3 [3.9 - 6.6] | 0.4 [−0.2 - 2.1] | 0.4 [0.0 - 2.0] | 0.1 [−0.7 - 0.9] | .02[d] | .007[d] | .68[d] |
| Total cholesterol (mmol/L), mean (SD) or median [IQR] | 4.9 (1.3) | 4.8 (1.2) | 5.0 (1.2) | 0.0 [−0.2 - 0.2] | 0.0 [−0.2 - 0.0] | 0.1 [−0.5 - 0.2] | >0.99[d] | .05[d] | .28[d] |
| LDL-C (mmol/L), mean (SD) or median [IQR][k] | 2.9 (1.1) | 2.9 (1.1) | 3.0 (1.0) | −0.1 [−0.3 - 0.1] | 0.0 [−0.2 - 0.0] | 0.0 [−0.3 - 0.3] | .18[d] | .04[d] | .48[d] |
| HDL-C (mmol/L), mean (SD) or median [IQR][l] | 1.11 (0.23) | 1.12 (0.19) | 1.1 (0.2) | −0.01 (0.12) | 0.0 [−0.1 - 0.0] | 0.0 [−0.1 - 0.1] | .56 | .27 | .55 |
| Triglycerides (mmol/L), median (IQR) | 1.88 [1.35 - 2.45] | 1.68 [1.21 - 1.90] | 1.8 [1.1 - 2.2] | 0.14 [0.00 - 0.47] | 0.0 [0.0 - 4.0] | 0.0 [−0.4 - 0.1] | **.003**[d] | .15[d] | .25[d] |
| hs-CRP (mg/L), median [IQR][m] | 3.0 [1.2 - 5.2] | 2.5 [1.1 - 3.9] | 2.4 [1.1 - 3.9] | 0.1 [-0.1 - 0.7] | 0.0 [0.0 - 0.8] | 0.1 [−0.3 - 1.0] | .14[d] | .03[d] | .27[d] |
| ALAT (IU/L), median [IQR][n] | 21.4 [18.2 - 30.2] | 23.2 [18.4 - 32.0] | 22.1 (18.7 - 29.1) | 0.0 [−6.8 - 2.8] | 0.0 [−5.0 - 2.9] | 0.8 [-1.6 - 7.0] | .37[d] | .80[d] | .18[d] |

| Characteristics | Baseline (n=38) | 12-week (n=38) | Month 9 (n=38) | Change from baseline to Week 12 | Change from baseline to Month 9 | Change from Week 12 to Month 9 | *P* value baseline versus Week 12[a] | *P* value baseline versus Month 9[a] | *P value* Week 12 versus Month 9[a] |
|---|---|---|---|---|---|---|---|---|---|
| ASAT, (IU/L), median (IQR)[o] | 20.8 [17.9 - 24.8] | 22.3 [18.0 - 25.5] | 19.4 (16.7 - 23.2) | 0.4 [−2.5 - 2.5] | 0.0 [−0.2 - 3.8] | 0.8 [−1.2 - 4.1] | .53[d] | .16[d] | .22[d] |
| Fibrosis-4 Index, median (IQR)[p] | 1.08 [0.78 - 1.34] | 1.08 [0.75 - 1.21] | 1.0 (0.7 - 1.2) | 0.01 [−0.06 - 0.07] | 0.0 [0.0 - 0.2] | 0.8 [−1.2 - 4.1] | .58[d] | .02[d] | **.03**[d] |

[a]Paired *t* tests were computed for approximately normally distributed data.

[b]N/A: not applicable.

[c]MRI-PDFF: magnetic resonance imaging proton density fat fraction.

[d]For nonnormal data, nonparametric tests were computed (Wilcoxon signed-rank tests).

[e]CAP: controlled attenuation parameter.

[f]Measured by dual-energy ray absorptiometry.

[g]$HbA_{1c}$: glycated hemoglobin A1c.

[h]s-glucose: Serum glucose.

[i]s-insulin: Serum insulin.

[j]HOMA-IR: homeostatic model assessment of insulin resistance.

[k]LDL-C: low-density lipoprotein cholesterol.

[l]HDL-C: high-density lipoprotein cholesterol.

[m]hs-CRP: high-sensitivity C-reactive protein.

[n]ALAT: alanine aminotransferase.

[o]ASAT: aspartate aminotransferase.

[p]Fibrosis-4: index for liver fibrosis.

At month 9, 29 out of 38 participants (76%) were classified as low risk of liver fibrosis based on the Fibrosis-4 Index, compared to 27 out of 39 (71%) at baseline. Six participants (16%) were classified as intermediate risk at month 9 compared to 7 (18%) at baseline, while 3 (8%) were classified as high risk at month 9 compared to 4 (11%) at baseline. Furthermore, the median Fibrosis-4 Index score significantly decreased over the 9-month study period, from 1.08 (IQR 0.78 - 1.34) at baseline to 1.0 (IQR 0.7 - 1.2) (*P*=.02).

No significant changes were observed in health-related quality of life (HRQoL; EQ-5D-5L), mental health (DASS-21), or medication adherence (MMAS-8) during the maintenance period, nor over the total study period (data not shown). In addition, the improvements in daily step count observed during the active 12-week program period were not sustained during the maintenance period, as recorded by the in-app step counter (data not shown).

### Associations Between App Engagement and Clinical Outcomes

An exploratory analysis was performed to assess the relationship between participants' in-app activity during the maintenance period and their clinical outcomes. Previously, it was reported that participants who were highly engaged with the app during the active 12-week period (defined as visiting the app at least 5 days per week) experienced greater weight loss and liver fat reduction compared to less engaged participants [12]. A similar pattern was observed during the maintenance period, with highly engaged participants having significantly greater weight loss and relative liver fat reduction than those with lower engagement levels (see Table S1 in Multimedia Appendix 2).

### Behavior Change

In an exploratory analysis, participants' self-reported dietary, exercise, and mental resilience behaviors were evaluated. Based on in-app questionnaires administered at Week 1 and Week 12, statistically significant improvements in healthy behaviors were found over the active period (see Table S2 in Multimedia Appendix 2).

### Adverse Events

In total, 26 adverse events were reported in the 6-month maintenance period (see Table S2 in Multimedia Appendix 2). No adverse events were considered related to the digital program as assessed by the investigator.

## Discussion

### Principal Findings

This study indicates that improvements in markers of cardiometabolic and liver-specific health markers, achieved during a 12-week active digital health program, can be maintained at 9 months, even without active coaching or new content being delivered during the 6-month maintenance period. We observed sustained weight loss, improvements in body composition, and reduction in liver fat, blood pressure, and glycemic control, all known as important key risk factors for cardiovascular disease [26]. At 9 months, we observed a significant reduction in hs-CRP levels and waist-to-hip ratio,

which was not present at 12 weeks. Both markers are indicators of cardiovascular disease risk. Elevated hs-CRP reflects low-grade systemic inflammation that plays a key role in the development of atherosclerosis, while an increase in waist-to-hip ratio is an indirect measure of abdominal obesity, another well-established cardiovascular disease (CVD) risk factor [27,28]. In addition, at 9 months, we observed significantly lower median Fibrosis-4 Index values compared to baseline. The Fibrosis-4 Index is a biomarker of liver fibrosis and can potentially be used as a noninvasive alternative to liver biopsy for diagnosis and managing liver disease [21]. Our 9-month data showed that more individuals were categorized in the low-risk fibrosis group and fewer individuals were categorized in the intermediate-risk and high-risk groups at 9 months compared to baseline, suggesting improvements in liver health over the 9-month study.

The improvements observed in this study are not only clinically significant but also have important public health implications. NAFLD, with its potential progression to more severe forms of liver disease, poses a substantial and increasing burden on health care systems globally as its prevalence continues to rise [29,30]. The positive and sustained health outcomes shown here suggest that scalable digital programs, such as SK-241, have the potential to alleviate this burden by providing effective and accessible solutions. This program can reduce the strain on health care professionals and ease the overall pressure on health care systems.

Previous studies on lifestyle and behavior change interventions among individuals with NAFLD demonstrate, in general, a low success rate in achieving long-term weight management, a high dropout rate, and poor adherence to the prescribed interventions [31-33]. Moreover, the evidence on the effectiveness of digital health interventions for NAFLD is limited. To the authors' knowledge, no previous studies have investigated the longer-term effectiveness of digital health programs for managing NAFLD [34]. Our study demonstrated that the improvements in weight loss, body composition, liver fat reduction, blood pressure control, insulin sensitivity, and glycemic control observed during the initial 12 weeks were sustained at 9 months. Suggesting that the program may induce potentially longer-lasting metabolic benefits. The digital nature of this program likely plays a role in fostering sustained engagement and support, helping participants adopt lasting behavior changes and effectively address the underlying causes of metabolic disturbances. Indeed, our exploratory analysis revealed that indicators of healthy behaviors, including diet, exercise, and mental resilience, were improved over the course of a 12-week program. These findings are consistent with the growing body of literature emphasizing the critical role of lifestyle modifications in mitigating metabolic dysfunction associated with NAFLD [35,36]. They contribute to the evolving narrative on the potential of digital health programs to serve as a scalable and accessible approach for improving cardiometabolic health in individuals with NAFLD and CVD in general.

Glucagon-like peptide-1 receptor agonists (GLP1-RAs) are considered promising treatment candidates for NAFLD and NASH and are indicated for the treatment of type 2 diabetes and obesity [37]. They have been shown to improve glycemic control and reduce weight, insulin resistance, and liver fat content [38]. A recent meta-analysis examining the effects of GLP1-RAs in individuals with NAFLD reported a mean weight loss of around 4 kg, comparable to the findings in our study. However, the mean relative reduction in liver fat content was greater, at 32% compared to the reduction observed in our study [38]. More recent studies have shown even higher weight loss and liver fat reductions after treatment with GLP1-RAs in individuals with NAFLD, making these medications a very promising treatment option for NAFLD [39]. Although these new medications have demonstrated strong clinical outcomes, there is potential for digital health solutions like SK-241 to complement GLP1-RAs treatments, provide additional support to patients using GLP1-RAs therapies, or serve as an alternative for those who cannot tolerate pharmacological treatment. However, this warrants further investigation.

The relatively high engagement and retention outcomes indicate that the 6-month maintenance period was well accepted by participants. However, we cannot assume that it was only or mainly the in-app activity during the maintenance period that drove the sustained health outcomes. Some may have continued to deploy the knowledge and behavior change tools they acquired in the first 12-week active program. Nevertheless, ongoing access to the app may serve as an important motivational tool for some participants, offering additional support to maintain health improvements and reinforce newly established habits. This suggests that the SK-241 program is a promising and valuable option in the comprehensive management of individuals with NAFLD and cardiometabolic conditions. However, further research is needed to determine the optimal duration of such a program to achieve sustained long-term benefits.

### Strengths and Limitations

A strength of this study was the high participant engagement and retention throughout the study period, enhancing the reliability of the longitudinal data. Furthermore, cardiovascular risk factors and liver fat content were assessed using objective and validated measures, with liver fat quantified using MRI-PDFF, a highly precise and reproducible imaging technique. Importantly, the observed reductions in liver fat content were accompanied by beneficial changes in cardiovascular risk markers. This temporal association supports the hypothesis that targeting liver fat may represent a feasible strategy to reduce cardiovascular risk, while also contributing to the ongoing scientific debate regarding the causal role of liver fat in cardiovascular disease [40,41].

Despite the encouraging findings, several limitations must be acknowledged. This was a preliminary, proof of concept study with a limited sample size of 28 participants and conducted under industry sponsorship. This characterization accurately reflects the early phase and exploratory nature of our work, which aimed to assess the feasibility and preliminary signals of efficacy, rather than to draw definitive clinical conclusions. The study was not powered to detect statistically significant differences in clinical endpoints or support subgroup analyses, such as by sex, which can be an important confounder in

NAFLD research [42]. The study's generalizability may be influenced by the specific characteristics of the study population, for example, all being Caucasian and with a relatively high education level. A potential limitation of this study is the absence of intermediate data collection points between the 12-week and 9-month follow-up assessments. While results at 9 months demonstrated sustained effects from the 12-week intervention period, the lack of more frequent measurements prevents a precise determination of whether a temporary "washout effect" (ie, a transient decline in the intervention's benefits) occurred during this interval. Future studies could benefit from more frequent data collection to track the trajectory of the intervention's effects more closely over time. Furthermore, missing values were imputed using baseline observation carried forward, which is a conservative approach but does not account for the fact that some individuals may have had different outcomes than their baseline values if they had attended the measurements, thus incurring a potential unmeasured bias. Further research is needed to explore the applicability of these results to a more diverse population and setting. In addition, the study did not address the potential impact of the program on other clinical outcomes, in addition to liver-related or cardiovascular morbidity and mortality, warranting further investigation. Moreover, as this was a single-arm study, results should be interpreted with caution. A full randomized controlled trial is needed to establish causal effects, to identify differences in outcomes by sex, and to validate these findings.

## Conclusions

This study adds to the growing evidence supporting the potential efficacy and sustainability of digital health programs in the management of NAFLD. Over a 6-month maintenance period, participants were able to sustain significant improvements in markers of liver and cardiometabolic health, indicating that the digital health program may produce long-lasting change, beyond the active phase. By targeting the root causes of metabolic disturbances and promoting long-term adoption of healthier habits, the SK-241 digital health programs provide holistic support for individuals with NAFLD. Moreover, the scalability and accessibility of the SK-241 digital health program have the potential to reduce the burden on health care systems. By empowering individuals to take an active role in managing their health remotely, these programs can alleviate pressure on traditional health care infrastructure and resources.

In conclusion, the findings of this study support the integration of digital health programs into the clinical management of NAFLD. Further research and implementation efforts are warranted to enhance the effectiveness and accessibility of these interventions, ultimately improving patient outcomes and reducing the burden on health care systems.

## Acknowledgments

Joanna McCarter provided medical writing support. Use of the MMAS-8 is protected by United States and international copyright laws. Permission for use is required. A Licensure agreement is available from: MMAR, LLC., www.moriskyscale.com. This study was sponsored and funded by Sidekick Health.

## Data Availability

The data sets generated or analyzed during this study are not publicly available due to restrictions in the informed consent form. Additional summary statistics will be provided upon reasonable request.

## Authors' Contributions

SB, HU, EFG, KS, TG, TK, SO, and VG contributed to conceptualization. SB, HU, EFG, KS, AI, BD, GEAM, TG, GB, SS, SO, and VG contributed to the methodology. SB, HU, EFG, AI, BD, SS, SO, and VG contributed to the investigation. SB and HU wrote the original draft. SB, HU, EFG, KS, AI, BD, GEAM, TG, TK, GB, SS, SO, and VG revised and edited the manuscript. SO acquired funding. SO and VG procured resources and supervised the study.

## Conflicts of Interest

HU, EFG, AI, BD, KS, TK, GEAM and TG are employed by Sidekick Health; SO is an employee and cofounder of Sidekick Health. SB received consultancy fees from Sidekick Health during the study period. SS, GB and VG have no competing interests to declare.

Multimedia Appendix 1
Supplemental report for MRI machine comparison.
[PDF File, 533 KB - cardio_v9i1e72074_app1.pdf ]

Multimedia Appendix 2
Supplementary tables for outcomes by in-app engagement, behavior change, and adverse events.
[DOCX File, 23 KB - cardio_v9i1e72074_app2.docx ]

## References


1. Riazi K, Raman MO, Taylor L, Swain MG, Shaheen AA. Dietary patterns and components in Nonalcoholic Fatty Liver Disease (NAFLD): what key messages can health care providers offer? Nutrients 2019 Nov 26;11(12):2878. [doi: 10.3390/nu11122878] [Medline: 31779112]
2. Mikolašević I, Filipec Kanižaj T, Targher G. Nonalcoholic fatty liver disease - a growing public health problem. Croat Med J 2021 Feb 28;62(1):1-3. [doi: 10.3325/cmj.2021.62.1] [Medline: 33660954]
3. Wong VS, Ekstedt M, Wong GH, Hagström H. Changing epidemiology, global trends and implications for outcomes of NAFLD. J Hepatol 2023 Sep;79(3):842-852. [doi: 10.1016/j.jhep.2023.04.036]
4. Adams LA, Anstee QM, Tilg H, Targher G. Non-alcoholic fatty liver disease and its relationship with cardiovascular disease and other extrahepatic diseases. Gut 2017 Jun;66(6):1138-1153. [doi: 10.1136/gutjnl-2017-313884] [Medline: 28314735]
5. Wainwright P, Byrne CD. Bidirectional relationships and disconnects between NAFLD and features of the metabolic syndrome. Int J Mol Sci 2016 Mar 11;17(3):367. [doi: 10.3390/ijms17030367] [Medline: 26978356]
6. Pipitone RM, Ciccioli C, Infantino G, et al. MAFLD: a multisystem disease. Therapeutic Advances in Endocrinology 2023 Jan;14. [doi: 10.1177/20420188221145549]
7. Konyn P, Ahmed A, Kim D. Causes and risk profiles of mortality among individuals with nonalcoholic fatty liver disease. Clin Mol Hepatol 2023 Feb;29(Suppl):S43-S57. [doi: 10.3350/cmh.2022.0351] [Medline: 36417893]
8. Satapathy SK, Bernstein DE, Roth NC. Liver transplantation in patients with non-alcoholic steatohepatitis and alcohol-related liver disease: the dust is yet to settle. Transl Gastroenterol Hepatol 2022;7:23. [doi: 10.21037/tgh-2020-15] [Medline: 35892055]
9. Tidwell J, Balassiano N, Shaikh A, Nassar M. Emerging therapeutic options for non-alcoholic fatty liver disease: A systematic review. World J Hepatol 2023 Aug 27;15(8):1001-1012. [doi: 10.4254/wjh.v15.i8.1001] [Medline: 37701920]
10. Ahmed IA, Mikail MA, Mustafa MR, Ibrahim M, Othman R. Lifestyle interventions for non-alcoholic fatty liver disease. Saudi J Biol Sci 2019 Nov;26(7):1519-1524. [doi: 10.1016/j.sjbs.2018.12.016] [Medline: 31762620]
11. Zhou R, Gu Y, Zhang B, et al. Digital therapeutics: emerging new therapy for nonalcoholic fatty liver disease. Clin Transl Gastroenterol 2023 Apr 1;14(4):e00575. [doi: 10.14309/ctg.0000000000000575] [Medline: 36854062]
12. Björnsdottir S, Ulfsdottir H, Gudmundsson EF, et al. User engagement, acceptability, and clinical markers in a digital health program for nonalcoholic fatty liver disease: prospective, single-arm feasibility study. JMIR Cardio 2024 Feb 15;8:e52576. [doi: 10.2196/52576] [Medline: 38152892]
13. Rinella ME, Lazarus JV, Ratziu V, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. J Hepatol 2023 Dec;79(6):1542-1556. [doi: 10.1016/j.jhep.2023.06.003] [Medline: 37364790]
14. Sanal MG, Gish RG, Méndez-Sánchez N, et al. NAFLD to MAFLD: collaboration, not confusion - rethinking the naming of fatty liver disease. Metab Target Organ Damage 2024;4(4):45. [doi: 10.20517/mtod.2024.54]
15. Byrne CD, Targher G. MASLD, MAFLD, or NAFLD criteria: have we re-created the confusion and acrimony surrounding metabolic syndrome? Metab Target Organ Damage 2024;4(2):10. [doi: 10.20517/mtod.2024.06]
16. Lonardo A, Bril F, Caldwell SH, et al. Researchers call for more flexible editorial conduct rather than abruptly adopting only the new MASLD nomenclature. J Hepatol 2024 May;80(5):e192-e194. [doi: 10.1016/j.jhep.2024.01.012] [Medline: 38336589]
17. Zhang X, Wong GH, Wong VS. Application of transient elastography in nonalcoholic fatty liver disease. Clin Mol Hepatol 2020 Apr;26(2):128-141. [doi: 10.3350/cmh.2019.0001n] [Medline: 31696690]
18. França M, Alberich-Bayarri Á, Martí-Bonmatí L, et al. Accurate simultaneous quantification of liver steatosis and iron overload in diffuse liver diseases with MRI. Abdom Radiol (NY) 2017 May;42(5):1434-1443. [doi: 10.1007/s00261-017-1048-0] [Medline: 28110367]
19. Hull H, He Q, Thornton J, et al. iDXA, Prodigy, and DPXL dual-energy x-ray absorptiometry whole-body scans: a cross-calibration study. J Clin Densitom 2009;12(1):95-102. [doi: 10.1016/j.jocd.2008.09.004] [Medline: 19028125]
20. Sumida Y, Yoneda M, Tokushige K, et al. FIB-4 first in the diagnostic algorithm of metabolic-dysfunction-associated fatty liver disease in the era of the global metabodemic. Life (Basel) 2021 Feb 14;11(2):143. [doi: 10.3390/life11020143] [Medline: 33672864]
21. Anstee QM, Berentzen TL, Nitze LM, et al. Prognostic utility of Fibrosis-4 Index for risk of subsequent liver and cardiovascular events, and all-cause mortality in individuals with obesity and/or type 2 diabetes: a longitudinal cohort study. Lancet Reg Health Eur 2024 Jan;36:100780. [doi: 10.1016/j.lanepe.2023.100780] [Medline: 38188279]
22. Henry JD, Crawford JR. The short‐form version of the Depression Anxiety Stress Scales (DASS‐21): Construct validity and normative data in a large non‐clinical sample. British J Clinic Psychol 2005 Jun;44(2):227-239. [doi: 10.1348/014466505X29657]
23. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008 May;10(5):348-354. [doi: 10.1111/j.1751-7176.2008.07572.x] [Medline: 18453793]
24. Berlowitz DR, Foy CG, Kazis LE, et al. Effect of intensive blood-pressure treatment on patient-reported outcomes. N Engl J Med 2017 Aug 24;377(8):733-744. [doi: 10.1056/NEJMoa1611179] [Medline: 28834483]

25. Bress AP, Bellows BK, King JB, et al. Cost-effectiveness of intensive versus standard blood-pressure control. N Engl J Med 2017 Aug 24;377(8):745-755. [doi: 10.1056/NEJMsa1616035] [Medline: 28834469]
26. Tamaki N, Kurosaki M, Takahashi Y, et al. Liver fibrosis and fatty liver as independent risk factors for cardiovascular disease. J of Gastro and Hepatol 2021 Oct;36(10):2960-2966. [doi: 10.1111/jgh.15589]
27. de Ferranti SD, Rifai N. C-reactive protein: a nontraditional serum marker of cardiovascular risk. Cardiovasc Pathol 2007 Jan;16(1):14-21. [doi: 10.1016/j.carpath.2006.04.006]
28. Cao Q, Yu S, Xiong W, et al. Waist-hip ratio as a predictor of myocardial infarction risk: A systematic review and meta-analysis. Medicine (Baltimore) 2018 Jul;97(30):e11639. [doi: 10.1097/MD.0000000000011639] [Medline: 30045310]
29. Fernando DH, Forbes JM, Angus PW, Herath CB. Development and progression of non-alcoholic fatty liver disease: the role of advanced glycation end products. Int J Mol Sci 2019 Oct 11;20(20):5037. [doi: 10.3390/ijms20205037] [Medline: 31614491]
30. Kumar R, Priyadarshi RN, Anand U. Non-alcoholic fatty liver disease: growing burden, adverse outcomes and associations. J Clin Transl Hepatol 2020 Mar 28;8(1):76-86. [doi: 10.14218/JCTH.2019.00051] [Medline: 32274348]
31. Vilar-Gomez E, Martinez-Perez Y, Calzadilla-Bertot L, et al. Weight loss through lifestyle modification significantly reduces features of nonalcoholic steatohepatitis. Gastroenterology 2015 Aug;149(2):367-378. [doi: 10.1053/j.gastro.2015.04.005] [Medline: 25865049]
32. Dudekula A, Rachakonda V, Shaik B, Behari J. Weight loss in nonalcoholic fatty liver disease patients in an ambulatory care setting is largely unsuccessful but correlates with frequency of clinic visits. PLoS ONE 2014;9(11):e111808. [doi: 10.1371/journal.pone.0111808] [Medline: 25375228]
33. Avery L, Exley C, McPherson S, Trenell MI, Anstee QM, Hallsworth K. Lifestyle behavior change in patients with nonalcoholic fatty liver disease: a qualitative study of clinical practice. Clin Gastroenterol Hepatol 2017 Dec;15(12):1968-1971. [doi: 10.1016/j.cgh.2017.06.011] [Medline: 28624648]
34. Seifi N, Bahari H, Soltani S, et al. The effects of electronic-based lifestyle interventions on nonalcoholic fatty liver disease: A systematic review. Digit HEALTH 2023;9:20552076231187597. [doi: 10.1177/20552076231187597] [Medline: 37529544]
35. Prabhakar O, Bhuvaneswari M. Role of diet and lifestyle modification in the management of nonalcoholic fatty liver disease and type 2 diabetes. Tzu Chi Med J 2021;33(2):135-145. [doi: 10.4103/tcmj.tcmj_86_20] [Medline: 33912410]
36. Stavropoulos K, Imprialos K, Pittaras A, Faselis C, Narayan P, Kokkinos P. Lifestyle modifications in non-alcoholic fatty liver disease and non- alcoholic steatohepatitis. CVP 2018 Mar 26;16(3):239-245. [doi: 10.2174/1570161115666170621080835]
37. Nauck MA, Quast DR, Wefers J, Meier JJ. GLP-1 receptor agonists in the treatment of type 2 diabetes - state-of-the-art. Mol Metab 2021 Apr;46:101102. [doi: 10.1016/j.molmet.2020.101102] [Medline: 33068776]
38. Mantovani AO, Petracca G, Beatrice G, Csermely A, Lonardo AO, Targher GO. Glucagon-like peptide-1 receptor agonists for treatment of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis: an updated meta-analysis of randomized controlled trials. Metabolites 2021;11(2):73. [doi: 10.3390/metabo11020073]
39. Nevola R, Epifani R, Imbriani S, et al. GLP-1 receptor agonists in non-alcoholic fatty liver disease: current evidence and future perspectives. Int J Mol Sci 2023 Jan 15;24(2):1703. [doi: 10.3390/ijms24021703] [Medline: 36675217]
40. Maurantonio M, Ballestri S, Odoardi MR, Lonardo A, Loria P. Treatment of atherogenic liver based on the pathogenesis of nonalcoholic fatty liver disease: a novel approach to reduce cardiovascular risk? Arch Med Res 2011 Jul;42(5):337-353. [doi: 10.1016/j.arcmed.2011.08.004] [Medline: 21843565]
41. Ahmed A, Cule M, Bell JD, Sattar N, Yaghootkar H. Differing genetic variants associated with liver fat and their contrasting relationships with cardiovascular diseases and cancer. J Hepatol 2024 Dec;81(6):921-929. [doi: 10.1016/j.jhep.2024.06.030] [Medline: 38960375]
42. Lonardo A, Nascimbeni F, Ballestri S, et al. Sex differences in nonalcoholic fatty liver disease: state of the art and identification of research gaps. Hepatology 2019 Oct;70(4):1457-1469. [doi: 10.1002/hep.30626] [Medline: 30924946]


## Abbreviations

**ALAT:** alanine aminotransferase
**ASAT:** aspartate aminotransferase
**CAP:** controlled attenuation parameter
**DASS-21:** Depression, Anxiety and Stress Scale
**DXA:** dual-energy x-ray absorptiometry
**EQ-5D-5L:** health-related quality of life
**GLP1-RA:** glucagon-like peptide-1 agonist
**$HbA_{1c}$:** glycated hemoglobin A1c
**HDL-C:** high-density lipoprotein-cholesterol
**HOMA-IR:** Homeostatic Model Assessment of Insulin Resistance
**hs-CRP:** high-sensitivity C-reactive protein
**LDL-C:** low-density lipoprotein-cholesterol

**MASLD:** Metabolic Dysfunction-Associated Steatotic Liver Disease
**MMAS-8:** Morisky Medication Adherence Scale
**MRI-PDFF:** magnetic resonance imaging-proton density fat fraction
**NAFLD:** nonalcoholic fatty liver disease
**NASH:** nonalcoholic steatohepatitis
**PRO:** Patient Reported Outcome
**T2D:** type 2 diabetes

*Edited by A Singh; submitted 17.02.25; peer-reviewed by A Lonardo, R Bipat; revised version received 11.07.25; accepted 05.08.25; published 12.09.25.*

*Please cite as:*
*Björnsdottir S, Ulfsdottir H, Gudmundsson EF, Dobies B, Sveinsdottir K, Isberg AP, Magnusdottir GEA, Gunnarsdottir T, Karlsdottir T, Bjornsdottir G, Sigurdsson S, Oddsson S, Gudnason V*
*Long-Term Feasibility and Outcomes of a Digital Health Program to Improve Liver Fat and Cardiometabolic Markers in Individuals With Nonalcoholic Fatty Liver Disease: Prospective Single-Arm Feasibility Study*
*JMIR Cardio 2025;9:e72074*
*URL: https://cardio.jmir.org/2025/1/e72074*
*doi:10.2196/72074*

# Analysis of a Medically Certified, Wrist-Worn Sensor for the Assessment of Heart Rate and Energy Expenditure During Daily Activities in Patients With Chronic Heart Failure or Coronary Artery Disease and Recreational Athletes: Validation Study

Ignace L J De Lathauwer[1,2], MD; Valerie A A van Es[3], MSc; Mayke M C J van Leunen[1,2], MD; Steven Onkelinx[4], PhD; Rutger W M Brouwers[2], MD, PhD; Danny A J P van de Sande[1], MD, PhD; Mathias Funk[2], PhD; Hareld M C Kemps[1,2], MD, PhD

[1]Department of Cardiology, Máxima Medisch Centrum, De Run 4600, Veldhoven, The Netherlands

[2]Department of Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands

[3]MoMiLab research unit, IMT School for Advanced Studies Lucca, Lucca, Italy

[4]Department of Allied Health Professions, Fontys University of Applied Sciences, Eindhoven, The Netherlands

**Corresponding Author:**
Ignace L J De Lathauwer, MD
Department of Cardiology, Máxima Medisch Centrum, De Run 4600, Veldhoven, The Netherlands

## Abstract

**Background:** Exercise capacity and lifestyle have proven to be important prognostic factors for cardiovascular patients. Both can be ameliorated through different preventive interventions. Cardiac rehabilitation and remote patient monitoring have been proven to reduce cardiac events and cardiovascular mortality. One of the most important goals of cardiac rehabilitation and remote patient monitoring is improving physical fitness and monitoring of cardiovascular parameters, which could predict cardiac deterioration. In order to monitor cardiac patients successfully, reliable and nonobtrusive devices to assess physical activity and cardiovascular parameters need to be available.

**Objective:** This validation study aims to determine the accuracy of the Philips Health Band (PHB), a noninvasive, wrist-worn, medically certified device, for the assessment of heart rate (HR) and energy expenditure (EE) in patients with chronic cardiovascular diseases and recreational athletes (RAs).

**Methods:** The assessment of HR and EE by the PHB was compared with indirect calorimetry (Oxycon Mobile [OM; CareFusion GmbH]) during an activity protocol consisting of daily activities. Three groups were assessed: patients with heart failure with reduced ejection fraction (HFrEF), patients with stable coronary artery disease (CAD) with preserved left ventricular ejection fraction, and RAs.

**Results:** A total of 57 patients were included: 19 with CAD, 19 with HFrEF, and 19 RAs. HR assessment in the HFrEF and CAD groups was significantly underestimated over the entire protocol by the PHB as compared to the OM, with poor and fair reliability, respectively. No significant difference in HR was found between the PHB and OM over the entire protocol for the RA group, with good reliability (HFrEF: mean difference 3.0; $P$<.001; intraclass correlation coefficient [ICC] 0.36; CAD: mean difference 2.7; $P$<.001; ICC 0.55; RA: mean difference 0.8; ICC 0.60). Assessment of EE showed an underestimation over the entire protocol for the RA and CAD group, with poor and fair reliability, respectively. The HFrEF group showed no significant difference in EE assessment over the entire protocol, with poor reliability (HFrEF: mean difference 0.09; ICC 0.32; CAD: mean difference 0.29; $P$<.001; ICC 0.46; RA: mean difference 0.79; $P$<.001; ICC 0.26). The responsiveness to detect within-patient changes in activity intensity of the PHB was moderate for the HFrEF and CAD groups and acceptable for the RA group.

**Conclusions:** HR and EE assessment of a medically certified noninvasive sensor using a photoplethysmogram and accelerometer showed poor accuracy and moderate responsiveness during an activity protocol reflecting daily living activities in patients with stable CAD and chronic HFrEF. Accuracy of HR in RAs was good and the responsiveness for both HR and EE was acceptable. This research confirms previous research and stresses the need for better patient-specific algorithms in noninvasive sensors, taking cardiovascular pathology and medication usage into account, for assessing HR and EE prior to their implemention in patient care.

***(JMIR Cardio 2025;9:e69343)*** doi:10.2196/69343

**KEYWORDS**

cardiac diseases; noninvasive device; smartwatch; validation studies; monitoring cardiac patients

## Introduction

Exercise capacity is known to be an important prognostic factor in patients with cardiovascular disease. Coronary artery disease (CAD) and heart failure (HF) are two of the most prevalent cardiovascular diseases, affecting millions of people worldwide [1]. In patients with CAD, regular moderate-intensity physical activity (PA) is associated with an increase in peak aerobic capacity and a reduction in all-cause mortality [2]. Research has shown a similar effect for patients with HF, indicating that decreased exercise capacity is linked to an increased risk of atrial arrhythmias, mortality, and hospitalizations due to HF exacerbations [3]. A study demonstrated that, even after adjusting for age, exercise capacity remains the strongest predictor for risk of death in both patients with cardiovascular disease and healthy individuals undergoing exercise testing. Due to this, exercise capacity is a more powerful predictor for mortality among men than other established risk factors of cardiovascular disease [4].

Over the years, several preventive interventions have emerged to enhance the prognosis of patients with cardiovascular disease by coaching as well as monitoring their health status [5]. First, exercise-based cardiac rehabilitation (CR) after an acute coronary syndrome is associated with a reduction of the risk of repeated cardiac events and cardiovascular mortality [6,7]. Despite the benefits of CR, participation rates remain low due to factors such as long distances to CR facilities and patient age [8]. Consequently, telerehabilitation has been proposed as an innovative solution. Second, research has demonstrated that remote patient monitoring (RPM) of patients with HF is effective in reducing mortality and HF-related hospitalizations [9-12]. Current RPM interventions use spot measurements of weight, blood pressure, and heart rate (HR) to monitor patients. In order to optimize RPM interventions and enable telerehabilitation, noninvasive sensors are needed for continuous monitoring of cardiovascular parameters.

It is essential for a monitoring device to be accurate and responsive if implemented in patient care. The accuracy of a device is defined as the closeness of agreement between the monitoring device measurement and the true value [13]. Responsiveness of a device is defined as its ability to detect within-patient changes of exercise intensity or cardiovascular parameters over time and is therefore highly important in patients with cardiovascular disease to monitor progression or their overall health status.

Previous trials investigating commercially available sensors in healthy individuals have shown mixed results and cannot be directly extrapolated to patients with cardiovascular disease due to differences in cardiac function and medication use, which affect chronotropic competence [14-17]. This emphasizes the need for validation studies in patients with cardiovascular disease. Herkert et al [18] demonstrated that 2 wrist-worn devices performed poorly in estimating energy expenditure (EE) and detecting within-patient changes during low-to-moderate exercise intensities in patients with HF and CAD. Similarly, a study evaluating the first generation Apple Watch, in patients with cardiovascular disease, found clinically acceptable HR accuracy during exercise, but an overestimation of EE [19]. A recent systematic review demonstrated that while Fitbit devices accurately measured step count and Apple Watch reliably measured HR, none of the tested devices accurately estimated EE, and most were not validated in patients with cardiovascular disease [20]. These findings suggest ongoing technical progress but emphasize the need for population-specific validation before such devices can be reliably used in patient care.

The aim of this validation trial is to investigate the accuracy and responsiveness of a medically certified wrist-worn sensor, the Philips Health Band (PHB), for the assessment of HR and EE in 3 patient populations: patients with HF with reduced ejection fraction (HFrEF), patients with stable CAD and preserved left ventricular ejection fraction (LVEF), and recreational athletes (RAs). If the PHB shows sufficient accuracy and responsiveness for measuring HR and EE, and thus PA levels, it could be implemented in clinical care (eg, telerehabilitation, secondary prevention, and RPM) to provide health care workers with continuous cardiovascular data and give patients insights into and promote their PA in daily life.

## Methods

### Study Population

Patients were included based on their diagnosis to form 3 patient groups: patients with HFrEF, patients with stable CAD and LVEF, and RAs who have visited a sports cardiologist before. Stable CAD is defined as the presence of angina pectoris caused by one or more coronary artery stenosis, which previously resulted in an acute coronary syndrome and required intervention (coronary artery bypass grafting, percutaneous coronary intervention, or medical therapy). The condition is considered stable when symptoms have remained unchanged in frequency, severity, and duration over time [21]. RAs were defined as men or women, >35 years of age, who perform sports at least 30 weeks a year, with a minimum of 2.5 hours of the same sport or 1.5 hours of different sports each week [22]. All 3 groups were analyzed separately. Patients were recruited via their cardiologist in the outpatient clinic of the Máxima Medical Center, the Netherlands. Eligible patients were contacted by the principal investigator, who provided verbal and written information about the validation study. Patients were excluded from the study if they had permanent atrial fibrillation, hemodynamically significant valvular disease, neurological or orthopedic conditions impairing physical exercise capacity, severe pulmonary disease impairing exercise capacity, peripheral vascular disease, or cognitive impairment. Patients had to be able to speak Dutch to be included.

### Protocol

Patients completed a laboratory activity protocol consisting of daily household activities reflecting real-life situations (cooking,

table cleaning, and vacuuming), walking on a treadmill, and cycling. All activities were low-to-moderate intensity. The activity protocol was based on 2 similar studies in patients with cardiovascular disease, where it appeared to be functional and feasible in these patient groups [18,23]. The protocol was adjusted based on the patient population. Activity intensities were the highest for RAs, since they are in good condition and used to sport at higher intensities, while they were lower for patients with CAD and the lowest for patients with HFrEF. Cycling was done on 3 different loads, while walking was done at 3 different speeds and incline angles, all depending on the different patient groups. The duration of the entire protocol was around one hour. An overview of the protocol is shown in Table 1. The protocol was performed at the physical therapy department in the Máxima Medical Center under the supervision of a medical doctor.

**Table .** Activity protocol.

| Activity type and activity | Duration (min) | Resting (min) |
|---|---|---|
| Sedentary activities | | |
| Sitting | 5 | —[a] |
| Standing | 2 | — |
| Household activities | | |
| Cooking | 3 | 1 |
| Cleaning | 3 | 1 |
| Vacuuming | 3 | 3 |
| Cycling (ergometer), load | | |
| HFrEF[b] 0 W; CAD[c] 0 W; RA[d] 0 W | 3 | 3 |
| HFrEF 25 W; CAD 40 W; RA 50 W | 3 | 3 |
| HFrEF 50 W; CAD 70 W; RA 100 W | 3 | 3 |
| Walking (treadmill), speed-incline | | |
| HFrEF 2 km/h; CAD 4 km/h; RA 4 km/h-5% | 3 | 3 |
| HFrEF 4 km/h; CAD 5.5 km/h; RA 5.5 km/h-5% | 3 | 3 |
| HFrEF 2 km/h–5%; CAD 4 km/h–5%; RA 4 km/h–10% | 3 | 3 |
| Stairs | | |
| Ascending | 1 | 1 |
| Descending | 1 | 1 |

[a]Not applicable.

[b]HFrEF: heart failure with reduced ejection fraction.

[c]CAD: coronary artery disease.

[d]RA: recreational athlete.

## Criterion Measure

A CareFusion Oxycon Mobile (OM) device was used during the entire protocol to measure breath-by-breath oxygen ($VO_2$) uptake and carbon dioxide ($VCO_2$) production. This is a lightweight mobile device consisting of a facemask with a gas analyzer and a 12-lead electrocardiogram (ECG) sensor. The 12-lead ECG sensor was attached to the gas analyzer unit and strapped on a backpack, worn by the patient. The OM was connected to a computer where real-time data was gathered. Gas and volume calibration and ambient conditions were verified before the start of the protocol. The OM provides a reliable criterion measure as it has been validated before by comparing it with the gold standard of EE measurements, the Douglas Bag [24].

## Device

The Philips Health Band (PHB) is a Conformité Européenne–marked medical class IIa, wrist-worn device that measures and tracks movement and physiological parameters of the wearer. The PHB consists of different sensors, including a photoplethysmography sensor, an altimeter, and a tri-axial accelerometer. HR can be assessed through the photoplethysmography signal, while EE is estimated by an algorithm including basal metabolic rate (based on the wearer's gender, age, height, and weight), activity, and HR. Patients wore the PHB on their nondominant wrist. The PHB was connected to the Philips Actigraphy Server System. The Philips Actigraphy Server System includes a mobile phone app and a Philips Health Suite Data Platform, where the data can be viewed and extracted by the authorized clinician. The Philips Actigraphy Server System was supplied with the most recent firmware updates.

### Data Analysis

Raw data from the breathing and HR analysis of the OM (sample rate 0.5 Hz) and the processed HR and EE data of the PHB (sample rate 0.0167 Hz) were exported and imported into a custom-made MATLAB (MATrix LABoratory; MathWorks) analysis program (R2023b [23.2.0.2409890]). The entire activity bounds were analyzed.

First, the EE was calculated from the OM breath-by-breath measurements using the Weir equation [25]:

$$EE=[(3.941\times VO2)+(1.11\times VCO2)]\times 1.1440$$

Outliers (eg, abrupt movements) in the HR and EE data were detected using a Hampel filter. Values exceeding 3 SDs from the median, calculated over the data point itself and up to 3 neighboring elements, were considered outliers and replaced with the median of this local window [26]. Afterward, the OM data were down-sampled to 0.0167 Hz to enable a correct comparison between the PHB data and the OM data. Then, the HR and EE data of the criterion measure (OM) and the device (PHB) were matched according to the timestamps corresponding to the activities of the protocol, as represented in Table 1, and were ready for comparison.

### Statistical Analysis

To achieve 80% power to detect an ICC of 0.75 (hypothesis 0), which is considered to indicate excellent agreement, a sample size of 19 participants per study group was calculated. This applies under the alternative hypothesis that an ICC of 0.4 (hypothesis 1) corresponds to poor agreement in the groups HFrEF, CAD, and RA.

Descriptive statistics were used to describe the population according to baseline clinical characteristics. Normality of the data was assessed by visual inspection of histograms and by interpreting skewness and kurtosis [27]. Between‐group differences (HFrEF vs CAD vs RA) were tested by one‐way ANOVA for continuous variables (age and LVEF), with Bonferroni-corrected post hoc comparisons where appropriate, and by Chi-square tests for categorical variables (sex and medication use). A 2-sided $P<.05$ was considered significant. The accuracy of the PHB was assessed by calculating the mean (SD), mean differences, and mean average percentage error (MAPE) in HR and EE obtained from the PHB compared with the criterion measure, the OM. These values were calculated per activity and over the entire protocol, including resting time.

One-sample *t* tests were performed using mean differences (between the PHB and the OM) compared with zero (hypothesis 0) to identify agreement between the PHB and the criterion measure within reasonable limits (set at a 10% error zone). In addition, Bland-Altman plots were created to illustrate the level of agreement between the estimated HR and EE, and the HR and EE from the criterion measure, with mean bias and 95% upper and lower limits of agreement (LoA). Data falling outside the LoA were inspected but did not meet any predefined exclusion criteria, such as extreme physiological values, poor signal quality, or documented device malfunctions. While there may be systematic errors under specific conditions (eg, high-intensity activities), these data were retained to ensure the analysis reflects the full range of real-world conditions encountered in the dataset.

To assess the reliability of the PHB for each activity and the entire protocol, the ICC using 2-way mixed models with absolute agreement was used. The ICC was considered poor below 0.4, fair between 0.4 and 0.59, good between 0.6 and 0.74, and excellent above 0.75 [28]. The responsiveness of the OM and PHB was assessed using a paired *t* test during cycling at different speeds and walking at different speeds and incline angles. All data analyses were performed using MATLAB (R2023b [23.2.0.2409890]).

### Ethical Considerations

Written informed consent was provided by all patients after they had received both oral and written information about the study. The validation study was approved by the local medical ethical committee of the Máxima Medical Center (institutional review board approval number NL79217.015.21) and was conducted in accordance with the Declaration of Helsinki. Patients did not receive any form of financial or material compensation for their participation.

## Results

### Patient Characteristics

A total of 57 patients were included and completed the activity protocol. The patients were equally divided into 3 groups: patients with HFrEF (n=19, mean age 69.5 years, SD 9.3 years), patients with CAD (n=19, mean age 63.7 years, SD 8.1 years), and RAs (n=19, mean age 58.8 years, SD 10.7 years). There was a significant difference in age across groups (ANOVA $P=.004$) and in LVEF (ANOVA $P<.001$), whereas gender distribution did not differ significantly ($\chi^2_2 P=.35$). Patients across all groups were predominantly male, except for one female in the HFrEF group and 2 in the CAD group. The majority of HFrEF and CAD patients were using drugs affecting HR (19/19 HFrEF patients, 100%; 18/19 CAD patients, 95%), compared to only 5 in the RA group (26%). Between‐group differences were significant for β-blocker use ($\chi^2_2 P<.001$) and amiodarone use ($\chi^2_2 P=.04$), but not for calcium channel blockers ($P=.11$) or ivabradine ($P=.36$). Additional patient characteristics can be found in Table 2.

**Table .** Patient characteristics.

| Variables | HF[a] (n=19) | CAD[b] (n=19) | RA[c] (n=19) | *P* value |
|---|---|---|---|---|
| Age (years), mean (SD) | 69.5 (9.3) | 63.7 (8.1) | 58.8 (10.7) | .004 |
| Sex, n/N (%) | | | | |
| Male | 18/19 (95) | 17/19 (90) | 19/19 (100) | .35 |
| Female | 1/19 (5) | 2/19 (11) | 0 | |
| LVEF[d] (%), mean (SD) | 37.7 (7.5) | 58.8 (6.5) | 61.8 (3.6) | <.001 |
| HF etiology, n/N (%) | | | | |
| iCMP[e] | 9/19 (47) | —[f] | — | — |
| Non-iCMP | 10/19 (53) | — | — | — |
| Medication, n/N (%) | | | | |
| Beta-blocker | 17/19 (90) | 11/19 (58) | 2/19 (11) | <.001 |
| Calcium channel blocker | 2/19 (11) | 7/19 (37) | 3/19 (16) | .11 |
| Amiodarone | 3/19 (16) | 0 | 0 | .04 |
| Ivabradine | 1/19 (5) | 0 | 0 | .36 |

[a]HF: heart failure.

[b]CAD: coronary artery disease.

[c]RA: recreational athlete.

[d]LVEF: left ventricular ejection fraction.

[e]iCMP: ischemic cardiomyopathy.

[f]Not applicable.

All data from the PHB of one RA was lost due to a synchronization problem and all data from the OM of one patient with HFrEF was lost due to technical problems, which resulted in the exclusion of these 2 patients from the validation analysis. Stair walking activities of 5 RAs, 5 patients with CAD, and 2 patients with HFrEF were excluded from the analysis due to OM measurement failure during that specific activity.

## Accuracy

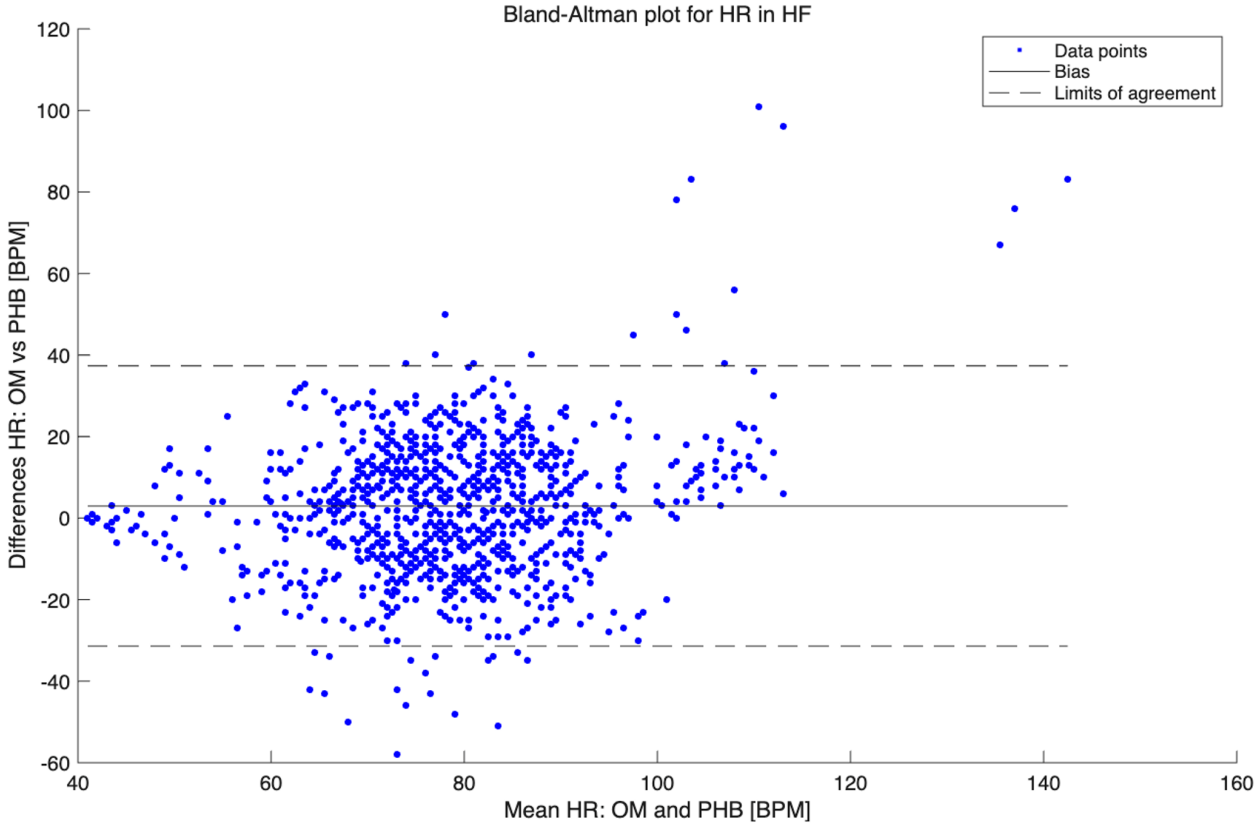
### *Patients With HFrEF*

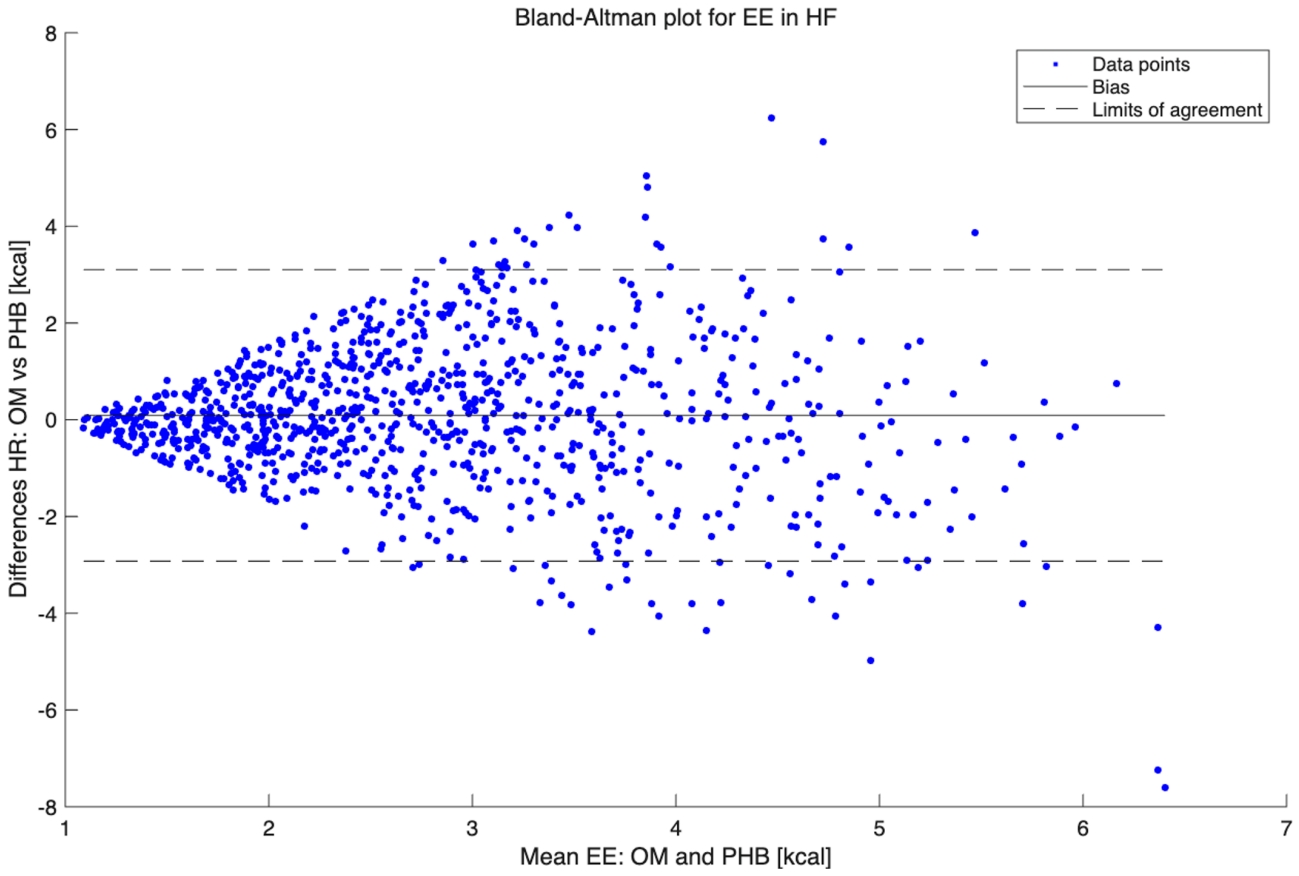
Table S1 in Multimedia Appendix 1 illustrates the accuracy of HR and EE measurements by the PHB for patients with HFrEF.

For HR, the mean (SD) over the entire protocol in the HFrEF group was 80.0 (16.9) beats per minute (bpm) for OM and 77.1 (13.6) bpm for PHB. The PHB significantly underestimated HR over the entire protocol, with a mean difference of 3.0 bpm ($P$<.001), showing a similar underestimation pattern for moderate-intensity household activities, cycling, and walking. For resting and low-intensity household activities, there were no significant differences between HR assessed by PHB and OM, except for standing, showing an underestimation of 3.2 bpm ($P$<.05). Bland-Altman plots for total HR measurements showed the PHB's underestimation, with wide LoA (lower LoA −30.7 bpm, upper LoA 36.7 bpm) (Figure 1). The bias was smaller for resting values (eg, for sitting, lower LoA −27.1 bpm, upper LoA 29.3 bpm) and increased with higher HR levels (eg, for cycling at 70 W, lower LoA −23.7 bpm, upper LoA 42.1 bpm). The ICCs for the total protocol indicated poor reliability, with a value of 0.36. The MAPE (SD) was 16.6 (13.9).

The EE results for the HFrEF group demonstrated a mean (SD) over the entire protocol of 2.86 (1.24) kcal for OM, and 2.76 (1.35) kcal for PHB (mean difference: 0.09 kcal, $P$=.06). However, significant underestimations were observed during climbing and walking down the stairs and cycling at 50 W, with mean differences of 0.54 kcal ($P$<.05), 1.04 kcal ($P$<.05), and 0.67 kcal ($P$<.001), respectively. It is important to note that resting and low-intensity household activities showed nonsignificant overestimations of EE, in contrast to other activities that were underestimated. Bland-Altman plots for total EE measurements indicated an underestimation by PHB, with a wide LoA for the total protocol (lower LoA −2.86 kcal, upper LoA 3.04 kcal; Figure 1). The bias for resting values was smaller (eg, for sitting, lower LoA −1.08 kcal, upper LoA 1.0 kcal), but increased toward EE values around 3 kcal, then stagnated (eg, cycling at 50 W, lower LoA −2.18, upper LoA 3.04). The ICCs for the total protocol indicated poor reliability, with a value of 0.32. The MAPE (SD) was 41.07 (40.53).

**Figure 1.** Bland-Altman plots heart rate and energy expenditure in patients with heart failure with reduced ejection fraction. BPM: beats per minute; EE: energy expenditure; HF: heart failure; HR: heart rate; OM: Oxycon Mobile; PHB: Philips Health Band.



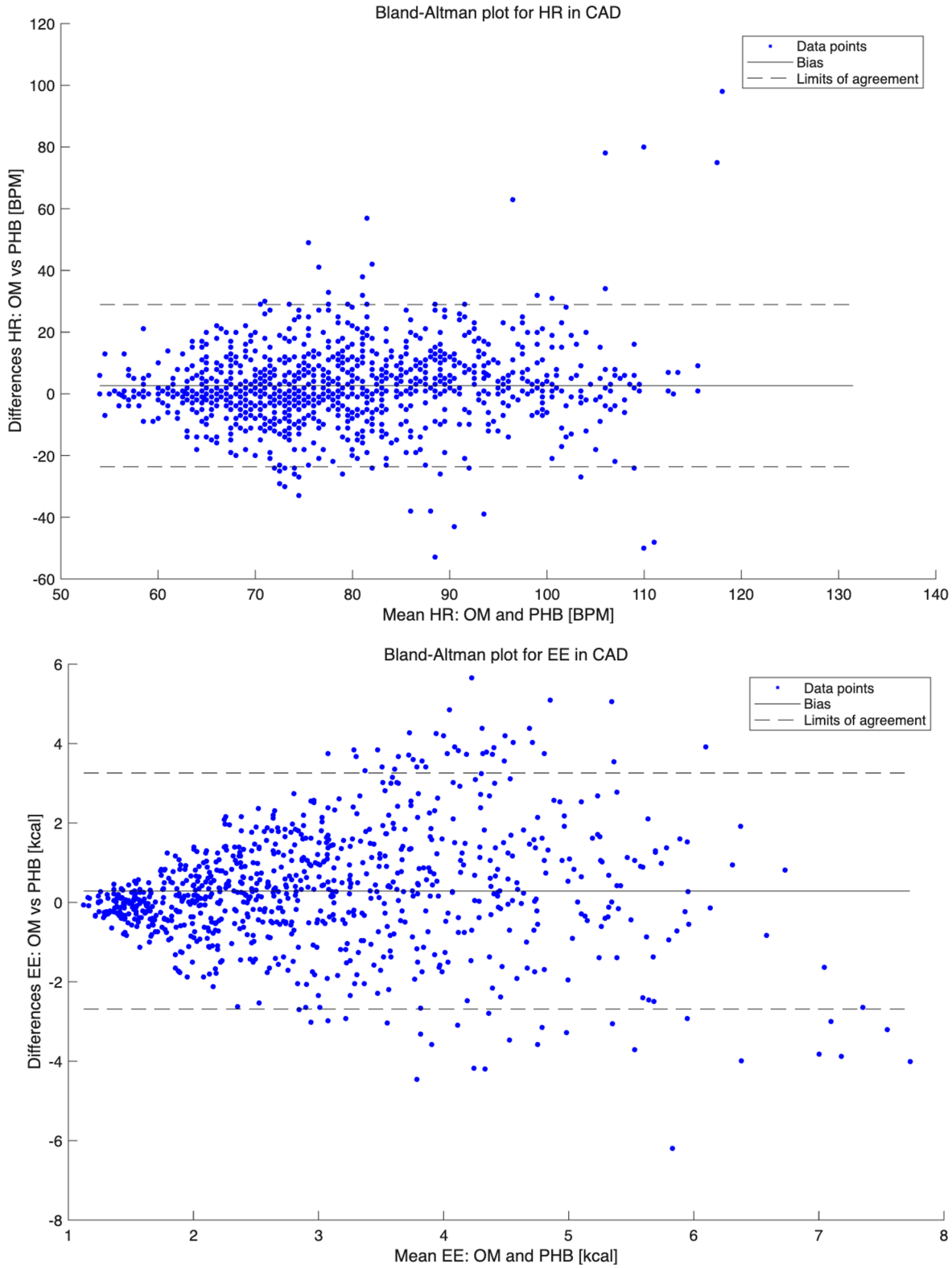



### *Patients With CAD*

Table S2 of Multimedia Appendix 1 demonstrates the accuracy of HR and EE measurements by the PHB for patients with CAD.

For HR, the mean (SD) over the entire protocol in the CAD group was 80.4 (15.2) bpm for OM and 77.7 (13.3) bpm for PHB. The PHB significantly underestimated HR over the entire protocol, with a mean difference of 2.7 bpm ($P$<.001), showing a similar underestimation pattern across all activities except for sitting, cleaning the table, and cycling at 0 W, with mean differences of 0.9 bpm, 1.7 bpm, and 0.8 bpm, respectively ($P$>.05). Bland-Altman plots for total HR measurements illustrated the PHB's underestimation, with a medium wide LoA (Figure 2). The PHB exhibited LoA from −23.1 bpm to 28.4 bpm. The bias was smaller for resting values (eg, for sitting, lower LoA −13.8 bpm, upper LoA 15.5 bpm) and increased with higher HR levels (eg, for cycling at 70 W, lower LoA −29.3 bpm, upper LoA 49.0 bpm). The ICCs for the total protocol indicated fair reliability, with a value of 0.55. The MAPE (SD) was 10.8 (10.7).

**Figure 2.** Bland-Altman plots for heart rate and energy expenditure in patients with coronary artery disease. BPM: beats per minute; CAD: coronary artery disease; EE: energy expenditure; HR: heart rate; OM: Oxycon Mobile; PHB: Philips Health Band.




The EE results for the CAD group demonstrated a mean (SD) over the entire protocol of 3.16 (1.48) kcal for OM and 2.88 (1.41) kcal for PHB. The PHB significantly underestimated EE across the entire protocol, with a mean difference of 0.29 kcal ($P$<.001). A similar underestimation pattern was observed for moderate-intensity household activities (except for climbing the stairs) and walking (except for walking at 4 km/h). For resting, lower intensity household activities, and cycling, the PHB showed nonsignificant differences compared to OM. Bland-Altman plots for total EE measurements indicated an underestimation by PHB, with wide LoA for higher EE values and narrow LoA for lower EE values (Figure 2). The PHB exhibited LoA from −2.63 kcal to 3.20 kcal for the total protocol. The bias for resting values was small (eg, for sitting,

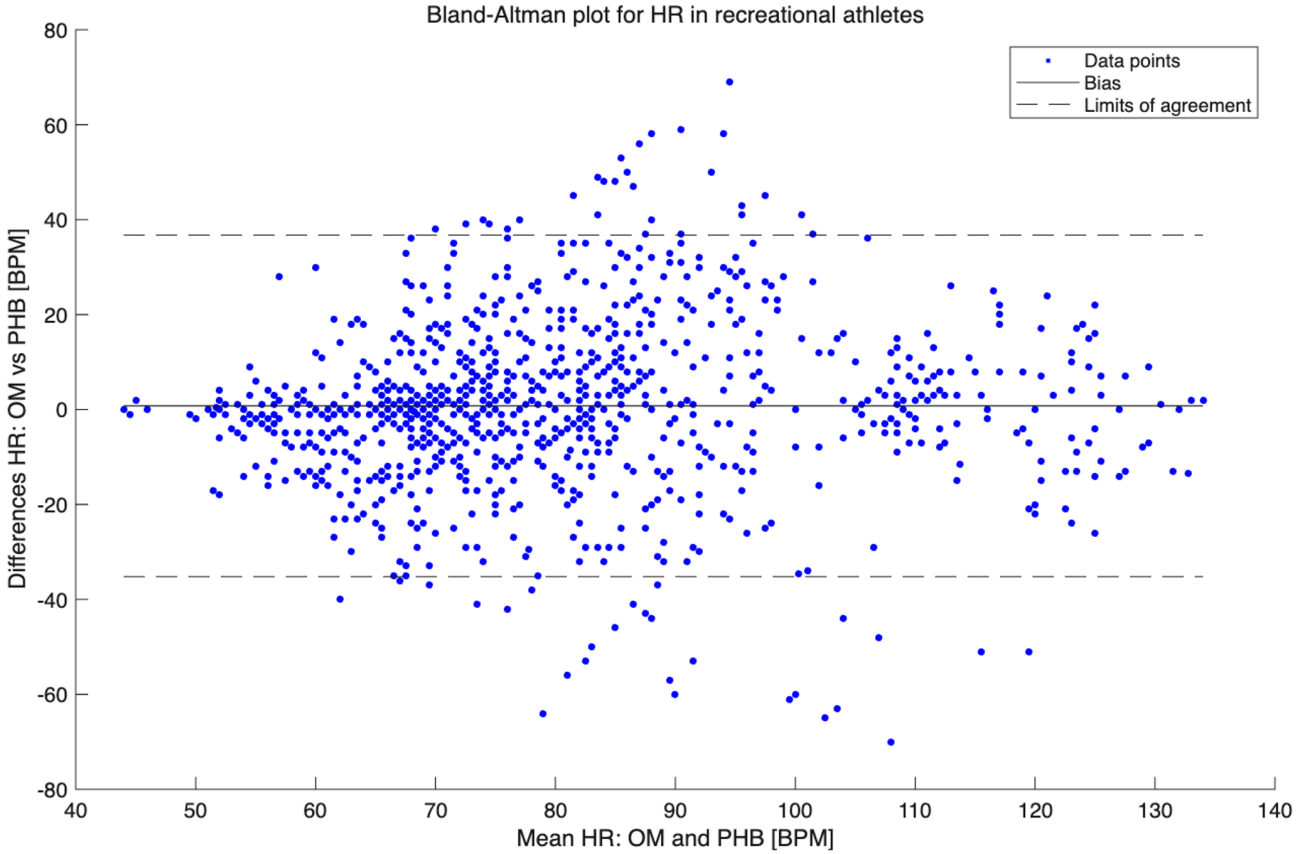
lower LoA −0.74 kcal, upper LoA 0.84 kcal) and increased with higher EE levels (eg, cycling at 70 W, lower LoA −3.71 kcal, upper LoA 4.09 kcal). The ICCs for the total protocol revealed fair reliability, with a value of 0.46. The MAPE (SD) was 35.66 (34.83).

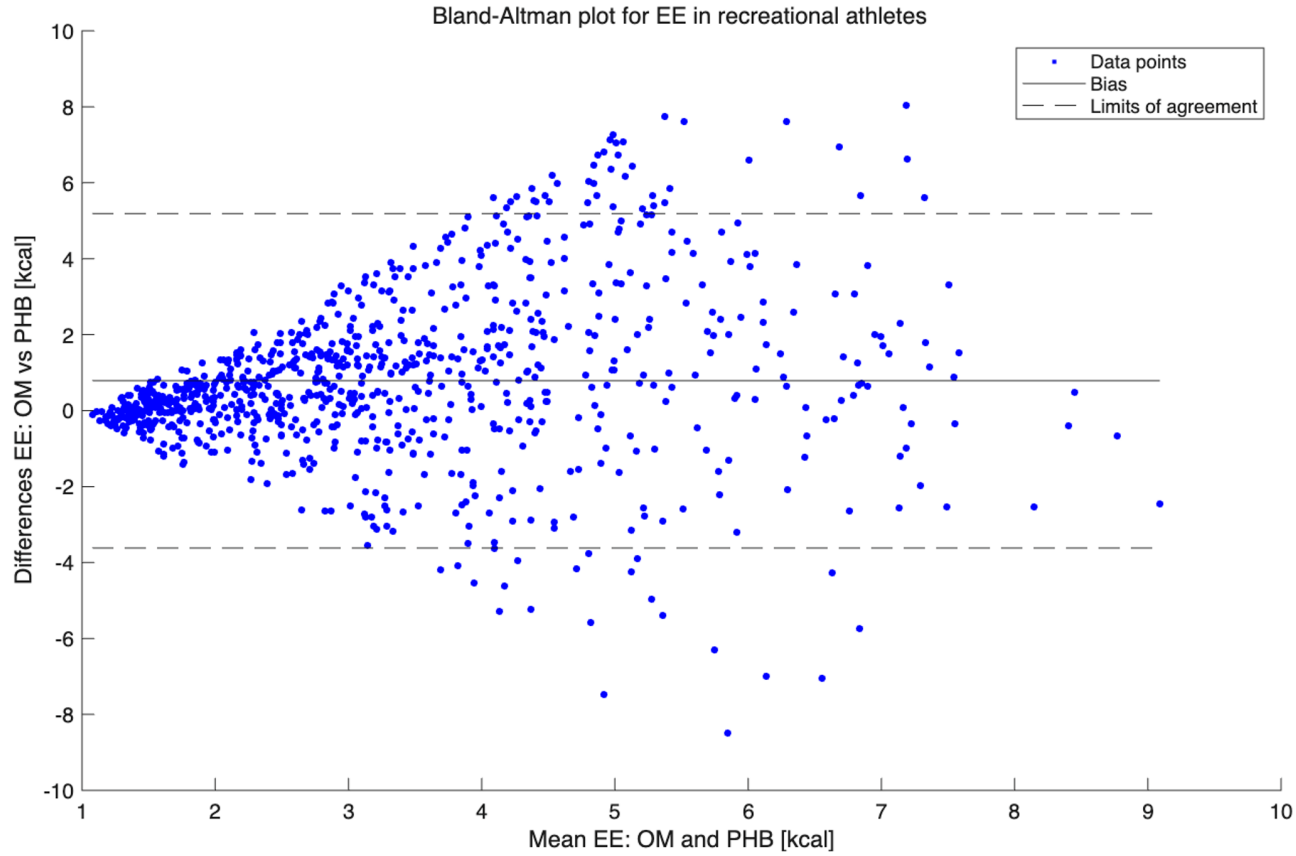
### *Recreational Athletes*

Table S3 of Multimedia Appendix 1 demonstrates the accuracy of HR and EE measurements by the PHB for RAs.

For HR, the mean (SD) over the entire protocol in the RA group was 81.0 (20.8) bpm for OM and 80.2 (19.5) bpm for PHB. The PHB showed nonsignificant ($P$>.05) underestimations over the entire protocol, with a mean difference of 0.8 bpm. Significant underestimations were found only for walking at all speeds, cycling (except at 0 W), and standing. For the other activities, there were nonsignificant ($P$>.05) differences between HR measurements by OM and PHB. Bland-Altman plots for total HR measurements illustrated the PHB's underestimation, with LoA from −34.5 bpm to 36.0 bpm (Figure 3). The bias was smaller for resting values (eg, for sitting, lower LoA −14.1 bpm, upper LoA 13.6 bpm) and increased with higher HR levels (eg, for cycling at 70 W, lower LoA −22.3 bpm, upper LoA 44.0 bpm). The ICCs for the total protocol indicated good reliability, with a value of 0.60. The MAPE (SD) was 16.2 (17.2).

**Figure 3.** Bland-Altman plots for heart rate and energy expenditure in recreational athletes. BPM: beats per minute; EE: energy expenditure; HR: heart rate; OM: Oxycon Mobile; PHB: Philips Health Band.




The EE results for RAs demonstrated a mean (SD) over the entire protocol of 3.80 (SD 2.11) kcal for OM and 2.96 (SD 1.71) kcal for PHB. The PHB significantly underestimated EE across the entire protocol, with a mean difference of 0.79 kcal

($P$<.001). This underestimation pattern was consistent across most activities, with only nonsignificant ($P$>.05) underestimations for standing (mean difference of 0.08 kcal), cooking (mean difference of 0.04 kcal), and cycling at 0 W (mean difference of 0.18 kcal). Bland-Altman plots for total EE measurements indicated an underestimation by PHB, with wide LoA for higher EE values and narrower LoA for lower EE values (Figure 3). The bias increased until EE expenditures were around 5 kcal and then decreased. The PHB exhibited LoA from −3.53 kcal to 5.10 kcal for the total protocol. The bias for resting values was small (eg, for sitting, lower LoA −0.93 kcal, upper LoA 1.55 kcal) and increased with higher EE levels (eg, cycling at 100 W, lower LoA −3.66 kcal, upper LoA 6.41 kcal). The ICCs for the total protocol revealed poor reliability, with a value of 0.26. The MAPE (SD) was 42.87 (38.51).

### Responsiveness

Table S4 of Multimedia Appendix 1 shows the ability of PHB to detect within-patient changes in cycling and walking activities.

#### *Patients With HFrEF*

For HR responses, the PHB was able to detect within-patient changes when cycling at 25 W versus 50 W (mean difference −1.28 bpm; $P$=.02) and when walking at 2 km/h versus 4 km/h (mean difference −3.78 bpm; $P$=.009). Note that differences in HR between cycling at 0 W versus 25 W were nonsignificant as measured by the OM. For EE responses, the PHB detected within-patient changes for cycling at 0 W versus 50 W (mean difference −0.44 kcal; $P$<0.001) and at 25 W versus 50 W (mean difference −0.42 kcal; $P$<.001). However, the PHB was not able to detect within-patient changes in EE for walking. It should be noted that differences in EE were nonsignificant for cycling at 0 W versus 25 W and for walking at 2 km/h versus 2 km/h with a 5% slope, as was measured by the OM.

#### *Patients With CAD*

For HR responses, the PHB was able to detect within-patient changes when cycling at 0 W versus 70 W (mean difference −4.32 bpm; $P$=.003) and when cycling at 40 W versus 70 W (mean difference −3.76 bpm; $P$<.001). For walking, the PHB was able to detect within-patient changes when walking at 4 km/h versus walking at 4 km/h with a 5% slope (mean difference −4.26 bpm; $P$<.001). Note that there were no significant differences between walking at 5.5 km/h versus walking at 4 km/h with a 5% slope, as measured by the OM. For EE responses, the PHB can detect within-patient changes when cycling at 0 W versus 70 W (mean difference −0.74 kcal; $P$<.001) and when cycling at 40 W versus 70 W (mean difference −0.57 kcal; $P$<.001). However, the PHB was not able to detect within-patient changes in EE for walking at different speeds and slopes. It should be noted that there were no significant changes in EE when walking at 5.5 km/h versus walking at 4 km/h with a 5% slope, as measured by the OM.

#### *Recreational Athletes*

For HR responses, the PHB was able to detect within-patient changes for cycling and walking at different watts, speeds, and slopes, except when cycling at 0 W versus 50 W. However, no significant differences were present when walking at 5.5 km/h with a 5% slope versus walking at 4 km/h with a 10% slope. For EE responses, the PHB detected within-patient changes for cycling and walking at different watts, speeds, and slopes, except for walking at 4 km/h with a 5% slope versus walking at 5.5 km/h with a 5% slope. There were no significant differences between walking at 5.5 km/h with a 5% slope versus walking at 4 km/h with a 10% slope, as measured by the OM.

## Discussion

### Principal Findings

This validation trial demonstrated poor accuracy of the PHB for monitoring HR in patients with HFrEF and patients with CAD, while there was no significant difference between the PHB and OM in the RA group, showing its ability to correctly measure HR in a healthier population. For all 3 groups, there was a pattern of underestimating HR and EE during more intense activities. EE was significantly underestimated in patients with CAD and RAs over the entire protocol. Responsiveness of the PHB demonstrated mixed results. The PHB was able to detect within-patient changes in HR and EE in RAs for almost all cycling loads and walking speeds. In patients with CAD and patients with HFrEF, the PHB demonstrated moderate to poor responsiveness to changes in cycling loads or walking speeds.

### Accuracy

Our study showed that the PHB demonstrates poor accuracy for measuring HR in patients with HFrEF and patients with CAD during moderate intensity activities. This is in contrast to previous studies investigating commercially available wrist-worn photoplethysmography sensors. Blok et al [29] investigated the accuracy of heartbeat detection using photoplethysmography sensors in patients with cardiovascular disease. They concluded that photoplethysmography sensors can determine HR with high accuracy in patients with cardiovascular disease. However, these measurements were made in the resting state. During activities, photoplethysmography signals are often contaminated by motion artifacts and noise, which deteriorate the signal quality and pose significant challenges on HR monitoring. This has led to different research suggesting algorithms for accurate HR tracking even in the presence of motion artifacts and noises [30]. Novel photoplethysmography-based sensors are integrated with algorithms for HR estimation even during activities. Kim et al [31] validated 2 new commercially available smartwatches for the assessment of HR during a cardiopulmonary exercise test in patients with CAD. They concluded that these newer devices show high concordance with the gold-standard ECG measurement. These results are also in contrast to the findings from our validation trial. A possible explanation for this might be the difference in the activity protocol. While Kim et al [31] validated the photoplethysmography sensors during a cardiopulmonary exercise test, we tried to validate the PHB sensor during an activity protocol with household activities reflecting real-life situations. These household activities included cooking, table cleaning, and vacuuming, which require more wrist movements. The placement of the photoplethysmography sensor on different body parts affects the severity of motion artifacts. Wrist placement is convenient since the photoplethysmography sensor can be integrated into

smartwatches and fitness trackers, but the wrist is more prone to motion artifacts and sensor detachment due to hand movement [32]. Moreover, the skin on the wrist moves more than other body parts, affecting sensor stability, likely influencing signal quality [32]. Conversely, placing the photoplethysmography sensor higher on the underarm or on the upper arm could reduce motion artifacts, since these areas experience less motion during daily activities, and skin movement is minimal compared to the wrist.

Another finding demonstrates a significant difference in the accuracy of HR between patients with HFrEF and patients with CAD compared to the . A possible explanation for this difference might be the patient's medication use and their cardiovascular pathology. Almost all patients in both HFrEF and CAD groups used drugs affecting their HR. In the RA group, only 26% of participants used drugs affecting HR. In addition, patients with HFrEF often endure chronotropic incompetence, which might affect HR estimation by the algorithm analyzing the photoplethysmography signal. The mechanism behind this in cardiovascular patients is that photoplethysmography-based HR measurement algorithms rely on detecting pulsatile blood volume changes in the peripheral microvasculature, which are often attenuated in these patients by reduced stroke volume and peripheral vasoconstriction. This leads to lower signal amplitude and distorted waveform morphology that can cause missed beats or misidentified peaks [33]. Chronotropic incompetence further narrows the dynamic range of HR changes, challenging algorithms tuned to larger beat-to-beat interval variability [33]. Blunting of the systolic upstroke by β-blockers and other rate-controlling medications alters the temporal features critical for peak detection. Moreover, other conditions like valvular heart disease or peripheral artery disease may similarly impair photoplethysmography signal quality due to altered vascular compliance or flow characteristics [34]. This stresses the need for more patient-specific algorithms for assessing HR through photoplethysmography signals. Potential pathways include dynamic peak detection thresholds, signal quality indexing, or machine-learning models trained on data from cardiovascular populations. These models could be tailored based on known patient characteristics (eg, presence of heart failure and medication use) and integrated into device firmware [35].

Our findings demonstrate a statistically significant difference in HR measurements between the PHB sensor and the OM in both HFrEF and CAD groups. However, the mean difference of approximately 3 bpm across the entire activity protocol is relatively small, raising questions about the clinical relevance of this discrepancy. Nonetheless, our data also demonstrates that the bias in HR measurements increases with rising HR values. This indicates that the measurement error becomes more pronounced at higher intensities, potentially leading to clinically meaningful discrepancies, particularly in contexts where the PHB sensor is used to support clinical decision-making, such as in RPM or cardiac telerehabilitation.

Our trial demonstrated that the PHB significantly underestimated EE over the entire protocol for patients with CAD and RAs. Gemini et al [20] conducted a systematic review examining studies that investigated the accuracy and acceptability of commercially available smartwatches. Of the 24 included studies, 22 assessed PA using EE as the outcome measure. Overall, all sensors demonstrated a MAPE of over 30%, indicating poor accuracy across all devices for assessing EE. The underestimation of EE by noninvasive sensors has also been observed in other studies. This aligns with our findings. All three groups showed an increase in underestimation with increasing activity intensity. This is in contrast to the findings from Herkert et al [18], who investigated 2 commercially available activity trackers in patients with CAD and patients with HFrEF. They observed an overestimation of EE over the entire protocol and an increase in overestimation when the activities intensify. This difference may possibly be explained by the variation in algorithms used to estimate EE. An alternative explanation for the underestimation of EE and its increase with intensified activities possibly lies within the HR sensor. Most algorithms for predicting EE in wrist-worn sensors are based on HR and accelerometer measurements. During this trial, we observed that the PHB significantly underestimated HR in patients with HFrEF and patients with CAD. Since these HR measurements are used to predict EE during these activities, it is expected that the underestimation would also be reflected in the EE prediction. Another explanation could be the simulation of the use of walking aids, which restrict arm movement in patients, by holding the handlebars of the treadmill. This restriction leads to decreased accelerometer measurements, resulting in a lower prediction of EE during those activities.

## Responsiveness

The PHB was able to detect some changes in both walking and cycling loads in patients with HFrEF and patients with CAD. However, the responsiveness of the PHB in RAs was a lot better compared to HFrEF and CAD groups. Research investigating the responsiveness of wrist-worn devices is scarce, especially in patients with cardiovascular disease; almost all trials focus their research solely on accuracy. Responsiveness is an important feature of smart devices for monitoring exercise activities at home. Herkert et al [18] investigated the responsiveness of 2 commercially available wrist-worn devices in patients with HFrEF and patients with CAD. They concluded that both sensors showed poor performance in detecting within-patient changes in the low-to-moderate exercise intensity domain. These findings are confirmed by our validation trial. Even though the PHB showed better responsiveness, there is still a lot of room for improvement, stressing the need for better algorithms for detecting within-patient changes during exercises for patients with cardiovascular disease.

## Future Perspectives

Our study clearly shows that even measurements of medically certified devices, using photoplethysmography and accelerometer to assess HR and EE, should be interpreted with caution for patients with cardiovascular disease. More studies with patients with cardiovascular disease and noninvasive sensors, using photoplethysmography and accelerometer, for assessing HR and EE should be done to enhance algorithm development. It is crucial that these trials extract raw photoplethysmography and accelerometer signals for better algorithm development. In addition, it is important that the validation of these new algorithms is conducted using an activity

protocol that reflects the patients' daily lives, rather than solely during exercise tests or rest measurements. Furthermore, future validation studies should not only focus on accuracy but also on the responsiveness of the sensors, as this is crucial for detecting within-patient changes throughout the day. Finally, to address existing barriers that hinder the usage of mHealth solutions and to assist health care professionals in evaluating the level of available evidence, a task force initiated by the European Society of Cardiology regulatory affairs committee formulated both general and specific criteria through a consensus process. These criteria should be consulted before considering the implementation of noninvasive devices in health care settings to ensure patient safety [36].

### Strengths and Limitations

A strength of this trial is that both patients with chronic cardiovascular diseases and RAs are included. In the results, there is a significant difference between the accuracy of the PHB for RA and for patients with cardiovascular disease in both accuracy and responsiveness. Stressing the need for algorithms that take into account both the cardiovascular pathology of the patients and medication usage. A limitation of this trial is the fact that patients were tested in a laboratory setting, even though the activity protocol consists of activities reflecting patients' daily life. This means that the results might not be able to be extrapolated to free-living conditions. In addition, a single activity protocol per patient group has been used to maximize the reproducibility of the study. Future validation trials could consider personalizing activity protocols, as cardiac patients show substantial variability in peak $VO_2$ and anaerobic threshold. Another limitation lies within the patient population. The majority of patients were men, making it possible that these results are not applicable to women.

### Conclusion

HR and EE assessment of a medically certified noninvasive sensor using a photoplethysmography and accelerometer showed poor accuracy and moderate responsiveness during an activity protocol reflecting daily living activities in chronic cardiac patients (HFrEF and CAD). High accuracy was obtained for HR in RA, while responsiveness was acceptable. This research confirms previous research and stresses the need for better patient-specific algorithms, taking cardiovascular pathology and medication usage into account, for assessing HR and EE.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
The appendix file contains the tables with the results from the statistical analysis.
[DOCX File, 34 KB - cardio_v9i1e69343_app1.docx ]

## References

1. Benjamin EJ, Muntner P, Alonso A, et al. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. Circulation 2019 Mar 5;139(10):e56-e528. [doi: 10.1161/CIR.0000000000000659] [Medline: 30700139]
2. Vooturi S, Anil P, Monica Y. Effects of exercise training and physical activity in patients with coronary artery disease. Indian J Clin Cardiol 2023 Jun;4(2):98-100. [doi: 10.1177/26324636231158460]
3. Sakai T, Yagishita A, Morise M, et al. Impact of exercise capacity on the long-term incidence of atrial arrhythmias in heart failure. Sci Rep 2021 Sep 21;11(1):18705. [doi: 10.1038/s41598-021-98172-9] [Medline: 34548557]
4. Myers J, Prakash M, Froelicher V, Do D, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. N Engl J Med 2002 Mar 14;346(11):793-801. [doi: 10.1056/NEJMoa011858] [Medline: 11893790]
5. Hansen D, Abreu A, Ambrosetti M, et al. Exercise intensity assessment and prescription in cardiovascular rehabilitation and beyond: why and how: a position statement from the secondary prevention and rehabilitation section of the European Association of Preventive Cardiology. Eur J Prev Cardiol 2022 Feb 19;29(1):230-245. [doi: 10.1093/eurjpc/zwab007] [Medline: 34077542]
6. Taylor RS, Brown A, Ebrahim S, et al. Exercise-based rehabilitation for patients with coronary heart disease: systematic review and meta-analysis of randomized controlled trials. Am J Med 2004 May 15;116(10):682-692. [doi: 10.1016/j.amjmed.2004.01.009] [Medline: 15121495]
7. European Association of Cardiovascular Prevention and Rehabilitation Committee for Science Guidelines, EACPR, Corrà U, et al. Secondary prevention through cardiac rehabilitation: physical activity counselling and exercise training: key components of the position paper from the Cardiac Rehabilitation Section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur Heart J 2010 Aug;31(16):1967-1974. [doi: 10.1093/eurheartj/ehq236] [Medline: 20643803]
8. van Engen-Verheul M, de Vries H, Kemps H, Kraaijenhagen R, de Keizer N, Peek N. Cardiac rehabilitation uptake and its determinants in the Netherlands. Eur J Prev Cardiol 2013 Apr;20(2):349-356. [doi: 10.1177/2047487312439497] [Medline: 22345699]
9. Lin MH, Yuan WL, Huang TC, Zhang HF, Mai JT, Wang JF. Clinical effectiveness of telemedicine for chronic heart failure: a systematic review and meta-analysis. J Investig Med 2017 Jun;65(5):899-911. [doi: 10.1136/jim-2016-000199] [Medline: 28330835]

10. Zhu Y, Gu X, Xu C. Effectiveness of telemedicine systems for adults with heart failure: a meta-analysis of randomized controlled trials. Heart Fail Rev 2020 Mar;25(2):231-243. [doi: 10.1007/s10741-019-09801-5] [Medline: 31197564]
11. Koehler F, Koehler K, Deckwart O, et al. Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): a randomised, controlled, parallel-group, unmasked trial. Lancet 2018 Sep 22;392(10152):1047-1057. [doi: 10.1016/S0140-6736(18)31880-4] [Medline: 30153985]
12. De Lathauwer ILJ, Nieuwenhuys WW, Hafkamp F, et al. Remote patient monitoring in heart failure: a comprehensive meta‐analysis of effective programme components for hospitalization and mortality reduction. Eur J Heart Fail 2025 Jan 20. [doi: 10.1002/ejhf.3568] [Medline: 39834044]
13. Menditto A, Patriarca M, Magnusson B. Understanding the meaning of accuracy, trueness and precision. Accred Qual Assur 2007 Jan 9;12(1):45-47. [doi: 10.1007/s00769-006-0191-z]
14. Wallen MP, Gomersall SR, Keating SE, Wisløff U, Coombes JS. Accuracy of heart rate watches: implications for weight management. PLoS One 2016;11(5):e0154420. [doi: 10.1371/journal.pone.0154420] [Medline: 27232714]
15. Chowdhury EA, Western MJ, Nightingale TE, Peacock OJ, Thompson D. Assessment of laboratory and daily energy expenditure estimates from consumer multi-sensor physical activity monitors. PLoS ONE 2017;12(2):e0171720. [doi: 10.1371/journal.pone.0171720] [Medline: 28234979]
16. Bai Y, Hibbing P, Mantis C, Welk GJ. Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. J Sports Sci 2018 Aug;36(15):1734-1741. [doi: 10.1080/02640414.2017.1412235] [Medline: 29210326]
17. Shcherbina A, Mattsson CM, Waggott D, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med 2017 May 24;7(2):3. [doi: 10.3390/jpm7020003] [Medline: 28538708]
18. Herkert C, Kraal JJ, van Loon EMA, van Hooff M, Kemps HMC. Usefulness of modern activity trackers for monitoring exercise behavior in chronic cardiac patients: validation study. JMIR mHealth uHealth 2019 Dec 19;7(12):e15045. [doi: 10.2196/15045] [Medline: 31855191]
19. Falter M, Budts W, Goetschalckx K, Cornelissen V, Buys R. Accuracy of Apple watch measurements for heart rate and energy expenditure in patients with cardiovascular disease: cross-sectional study. JMIR mHealth uHealth 2019 Mar 19;7(3):e11889. [doi: 10.2196/11889] [Medline: 30888332]
20. Germini F, Noronha N, Borg Debono V, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: systematic review of the literature. J Med Internet Res 2022 Jan 21;24(1):e30791. [doi: 10.2196/30791] [Medline: 35060915]
21. Task Force Members, Montalescot G, Sechtem U, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J 2013 Oct;34(38):2949-3003. [doi: 10.1093/eurheartj/eht296] [Medline: 23996286]
22. Van Middelkoop M, Kolkman J, Van Ochten J, Bierma-Zeinstra SMA, Koes BW. Risk factors for lower extremity injuries among male marathon runners. Scand J Med Sci Sports 2008 Dec;18(6):691-697. [doi: 10.1111/j.1600-0838.2007.00768.x] [Medline: 18266787]
23. Kraal JJ, Sartor F, Papini G, et al. Energy expenditure estimation in beta-blocker-medicated cardiac patients by combining heart rate and body movement data. Eur J Prev Cardiol 2016 Nov;23(16):1734-1742. [doi: 10.1177/2047487316667786] [Medline: 27625154]
24. Rosdahl H, Gullstrand L, Salier-Eriksson J, Johansson P, Schantz P. Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method. Eur J Appl Physiol 2010 May;109(2):159-171. [doi: 10.1007/s00421-009-1326-9] [Medline: 20043228]
25. WEIR J. New methods for calculating metabolic rate with special reference to protein metabolism. J Physiol 1949 Aug;109(1-2):1-9. [doi: 10.1113/jphysiol.1949.sp004363] [Medline: 15394301]
26. Pearson RK. Outliers in process modeling and identification. IEEE Trans Contr Syst Technol 2002;10(1):55-63. [doi: 10.1109/87.974338]
27. Kim HY. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod 2013 Feb;38(1):52-54. [doi: 10.5395/rde.2013.38.1.52] [Medline: 23495371]
28. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6(4):284-290. [doi: 10.1037/1040-3590.6.4.284]
29. Blok S, Piek MA, Tulevski II, Somsen GA, Winter MM. The accuracy of heartbeat detection using photoplethysmography technology in cardiac patients. J Electrocardiol 2021;67:148-157. [doi: 10.1016/j.jelectrocard.2021.06.009] [Medline: 34256184]
30. Zhu L, Kan C, Du Y, Du D. Heart rate monitoring during physical exercise from photoplethysmography using neural network. IEEE Sens Lett 2018;3(1):1-4. [doi: 10.1109/LSENS.2018.2878207]
31. Kim C, Song JH, Kim SH. Validation of wearable digital devices for heart rate measurement during exercise test in patients with coronary artery disease. Ann Rehabil Med 2023 Aug;47(4):261-271. [doi: 10.5535/arm.23019] [Medline: 37536665]
32. Fine J, Branan KL, Rodriguez AJ, et al. Sources of inaccuracy in photoplethysmography for continuous cardiovascular monitoring. Biosensors (Basel) 2021 Apr 16;11(4):126. [doi: 10.3390/bios11040126] [Medline: 33923469]
33. Mejía-Mejía E, Allen J, Budidha K, El-Hajj C, Kyriacou PA, Charlton PH. Photoplethysmography signal processing and synthesis. In: Kyriacou PA, Budidha K, editors. Photoplethysmography: Technology, Signal Analysis, and Applications: Academic Press; 2021:139-175.

34. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 2007 Mar;28(3):R1-39. [doi: 10.1088/0967-3334/28/3/R01] [Medline: 17322588]
35. Zhang Z, Pi Z, Liu B. TROIKA: a general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Trans Biomed Eng 2015 Feb;62(2):522-531. [doi: 10.1109/TBME.2014.2359372] [Medline: 25252274]
36. Caiani EG, Kemps H, Hoogendoorn P, et al. Standardized assessment of evidence supporting the adoption of mobile health solutions: a clinical consensus statement of the ESC Regulatory Affairs Committee. Eur Heart J Digit Health 2024 Sep 23;5(5):509-523. [doi: 10.1093/ehjdh/ztae042]

## Abbreviations

**bpm:** beats per minute
**CAD:** coronary artery disease
**CR:** cardiac rehabilitation
**ECG:** electrocardiogram
**EE:** energy expenditure
**HFrEF:** heart failure with reduced ejection fraction
**HR:** heart rate
**LoA:** limits of agreement
**LVEF:** left ventricular ejection fraction
**MAPE:** mean average percentage error
**OM:** Oxycon Mobile
**PA:** physical activity
**PHB:** Philips Health Band
**RA:** recreational athlete
**RPM:** remote patient monitoring

*Edited by N Ainani; submitted 27.11.24; peer-reviewed by N Scholte, T Kaihara; revised version received 09.07.25; accepted 06.08.25; published 30.09.25.*

*Please cite as:*
*De Lathauwer ILJ, van Es VAA, van Leunen MMCJ, Onkelinx S, Brouwers RWM, van de Sande DAJP, Funk M, Kemps HMC Analysis of a Medically Certified, Wrist-Worn Sensor for the Assessment of Heart Rate and Energy Expenditure During Daily Activities in Patients With Chronic Heart Failure or Coronary Artery Disease and Recreational Athletes: Validation Study*
*JMIR Cardio 2025;9:e69343*
*URL: https://cardio.jmir.org/2025/1/e69343*
*doi:10.2196/69343*

# Correction: Results of a Digital Multimodal Motivational and Educational Program as Follow-Up Care for Former Cardiac Rehabilitation Patients: Randomized Controlled Trial

Maxi Pia Bretschneider[1]; Wolfgang Mayer-Berger[2], Dr med; Jens Weine[3]; Lena Roth[1], MSc; Peter E H Schwarz[1], Prof Dr Med, hablil; Franz Petermann[2†], Prof Dr Med

[1]Department for Prevention and Care of Diabetes, Department of Medicine III, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, Dresden, Germany

[2]Klinik Roderbirken der Deutschen Rentenversicherung Rheinland, Leichlingen, Germany

[3]Vision2B GmbH, Erfurt, Germany

[†]deceased

**Corresponding Author:**
Lena Roth, MSc
Department for Prevention and Care of Diabetes, Department of Medicine III, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, Dresden, Germany

**Related Article:**

Correction of: https://cardio.jmir.org/2024/1/e57960

***(JMIR Cardio 2025;9:e73890)*** doi:10.2196/73890

In "Results of a Digital Multimodal Motivational and Educational Program as Follow-Up Care for Former Cardiac Rehabilitation Patients: Randomized Controlled Trial" (JMIR Cardio 2024;8:e57960) the authors noted two errors.

In the Results section of the Abstract, the following sentence:

> *Positive effects on secondary outcomes like body weight, blood pressure, and number of smokers only showed time effects, indicating no difference between the groups.*

Has been revised to:

> *Secondary outcomes like the body weight and cholesterol levels were significantly reduced in the intervention group, also in comparison with the control group.*

In addition, the degree for author Maxi Pia Bretschneider was removed as it was reported erroneously.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

*Submitted 13.03.25; this is a non–peer-reviewed article;accepted 14.03.25; published 12.05.25.*

*Please cite as:*
*Bretschneider MP, Mayer-Berger W, Weine J, Roth L, Schwarz PEH, Petermann F*
*Correction: Results of a Digital Multimodal Motivational and Educational Program as Follow-Up Care for Former Cardiac Rehabilitation Patients: Randomized Controlled Trial*
*JMIR Cardio 2025;9:e73890*
*URL: https://cardio.jmir.org/2025/1/e73890*
*doi:10.2196/73890*

# Predicting Atrial Fibrillation Ablation Outcomes: Machine Learning Model Development and Validation Using a Large Administrative Claims Database

Yijun Liu[1], BS; Mustapha Oloko-Oba[2], PhD; Kathryn A Wood[2], RN, PhD; Michael S Lloyd[3], MD; Joyce C Ho[4], PhD; Vicki Stover Hertzberg[2], PhD

[1]Department of Data and Decision Sciences, Emory University, 36 Eagle Row, Atlanta, GA, United States

[2]Nell Hodgson Woodruff School of Nursing, Emory University, 1520 Clifton Road NE, Atlanta, GA, United States

[3]Division of Cardiology, Department of Medicine, Emory University, 100 Woodruff Circle, Atlanta, GA, United States

[4]Department of Computer Science, Emory University, 400 Dowman Drive, Atlanta, GA, United States

**Corresponding Author:**
Joyce C Ho, PhD
Department of Computer Science, Emory University, 400 Dowman Drive, Atlanta, GA, United States

## Abstract

**Background:** Atrial fibrillation (AF) ablation is an effective treatment for reducing episodes and improving quality of life in patients with AF. However, long-term AF-free rates after AF ablation are inconsistent across the population, ranging from 50% to 75%. Patient selection relies on individual clinical assessment, highlighting a critical gap in population-level predictive analytics. While existing risk scores (eg, $CHADS_2$ [congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, and stroke], $CHA_2DS_2$-VASc [congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age, and sex category], CAAP-AF [coronary artery disease, left atrial diameter, age, AF, antiarrhythmic drugs, and female sex category]) have been applied to predict AF ablation outcomes, their performance in administrative claims data remains unclear. Leveraging large administrative claims databases represents an opportunity to develop standardized, scalable prediction models that could inform population health management and resource allocation at a national level.

**Objective:** This study utilizes machine learning (ML) models on claims data to explore if integrating *International Classification of Diseases* (*ICD*) billing codes outperforms traditional stroke and AF risk scores in predicting 1-year AF ablation outcomes.

**Methods:** We analyzed claims data from the Merative MarketScan Research Medicare database (2013‐2020) to identify 14,521 patients who underwent AF ablation. To predict 1-year AF-free outcomes, we developed logistic regression and extreme gradient boosting (XGBoost) models using demographic characteristics, comorbidity indices, and *ICD* diagnostic codes from the 2 years preceding ablation. Model predictions were compared with claims-based implementations of established risk scores—$CHADS_2$, $CHA_2DS_2$-VASc, and a modified CAAP-AF (without left atrial diameter and the number of failed antiarrhythmic drugs). The ML models were also assessed on subgroups of patients with paroxysmal AF, persistent AF, and both AF and atrial flutter from October 2015 onward.

**Results:** Among 14,521 patients (mean age 71.5, SD 5.31 y; n=5800, 39.94% female), AF ablation success occurred in 54.01% (n=7843). XGBoost achieved areas under the receiver operating characteristic curve (AUCs) of 0.528, 0.521, and 0.529 for the whole, female, and male AF ablation groups, respectively, and better discrimination than $CHADS_2$, $CHA_2DS_2$-VASc, and the modified CAAP-AF in all AF ablation groups (whole population, female, and male). While $CHA_2DS_2$-VASc and the modified CAAP-AF showed higher recall (>0.798), their precision (<0.540) was lower than XGBoost (0.552‐0.556). In subgroup analyses of *International Classification of Disease, Tenth Revision* (*ICD-10*) patients (n=7646), the models incorporating *ICD* codes demonstrated better performance than those using only demographic and comorbidity data across most AF subtypes, with the highest AUC (0.544) observed in patients with paroxysmal AF.

**Conclusions:** While the ML models achieved statistically significant improvements over claim-based implementations of established clinical risk scores (AUC 0.528‐0.544 vs 0.498‐0.505), the modest predictive performance highlights challenges in predicting procedural outcomes using administrative data that lack key clinical variables (eg, left atrial size and medication details). Our findings establish that while standardized outcome prediction using nationally available administrative data is technically feasible, current performance is insufficient for clinical decision-making and better suited for health system quality monitoring and comparative effectiveness research applications.

***(JMIR Cardio 2025;9:e77380)*** doi:10.2196/77380

**KEYWORDS**

administrative claims data; atrial fibrillation; atrial fibrillation ablation; machine learning; XGBoost; extreme gradient boosting

## Introduction

Although there is currently no cure for atrial fibrillation (AF), a major public health concern in the United States, AF ablation is the most effective treatment to restore normal sinus rhythm and decrease symptoms in episodes of paroxysmal or persistent AF, thereby reducing AF burden and improving quality of life [1-3]. AF is associated with an increased risk of cardiovascular events that may affect treatment outcomes. While various clinical risk factors are well understood, existing risk scores have shown inconsistent effectiveness in predicting AF ablation outcomes.

Existing risk scores, such as $CHADS_2$ (congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, and stroke) and $CHA_2DS_2$-VASc (congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age, and sex category), have traditionally been applied to predict stroke risk and are now also utilized in predicting outcomes following COVID-19, heart surgery, and AF ablation [4-11]. One risk prediction scale specifically designed to predict outcomes from AF ablation, the CAAP-AF (coronary artery disease, left atrial diameter, age, AF, antiarrhythmic drugs, and female sex category) risk score, addresses the presence or absence of coronary artery disease, the left atrial diameter, the presence of persistent AF, the number of antiarrhythmic drugs that have failed, and female sex [12].

Success rates for AF ablation from the literature vary based on individual clinical variables, such as the type of AF, left atrial size, or volume index [1,3,12,13], yet these variables are often difficult to access in large electronic health record (EHR) datasets. Patients can continue to experience episodes of AF following initial AF ablation with long-term AF-free rates after de novo AF ablation reported as 50%-75% [1,3,14]. Additionally, the chances of developing any complications after AF ablation range around 6%, with 0.1%-0.9% of patients experiencing complications that could result in death [15-17]. Given the modest success rates of AF ablation, the prediction of outcomes could be personalized to more easily identify those who would be most likely to benefit from AF ablation.

Machine learning (ML) has emerged as a powerful approach that leverages increased computational power with large datasets to help achieve complex decisions to guide clinical practice [18]. Artificial intelligence and ML have been used in the field of electrophysiology since the 1970s for automated electrocardiogram interpretation [18,19]. More recently, innovations in algorithms, development and labeling of large databases, and improvements in hardware and software have rapidly increased the role of ML in cardiac electrophysiology and cardiovascular imaging to identify predictors of patient outcomes [20]. Recent studies have demonstrated the potential for ML approaches in cardiovascular medicine, from achieving impressive diagnostic performance using novel data sources such as mobile phone acoustics for heart failure detection [21] to identifying practice gaps in stroke care guidelines [22] and showing predictive accuracy across various cardiovascular disease predictions [23]. ML has already been used to improve the prediction of AF ablation outcomes, primarily via EHRs. Nevertheless, health systems are not widely interoperable [24]; thus, extending these prognostic tools across multiple health systems is both costly and challenging. Studies utilizing EHR data have often been limited to datasets from 1 to 2 hospitals, limiting the generalizability of the models and hindering broad adoption [25,26].

Health insurance claims data, in comparison, are commonly collected, more readily available, and usually collected on a large national scale [27]. Although EHR data, which can include medications, laboratory data, and radiology reports, are more granular than claims data and can offer more accurate predictions, claims data's breadth and consistency across health systems can potentially provide stronger external validity [28] and more cost-effective scaling. A recent study applied ML models on health insurance data for cardiovascular outcome prediction and achieved area under the receiver operating characteristic curve (AUC) of 0.68 - 0.69 for heart failure readmission prediction [29], illustrating the potential for population-level insights using administrative databases. This wider coverage across patient populations and care settings may yield models that generalize more effectively, reducing the need for labor-intensive data extraction and curation that often blocks EHR-based projects. Furthermore, claims-based prognostic models can be used to enhance health care resource allocation by reducing unnecessary procedures in patients unlikely to benefit and increasing access to this effective therapy for appropriate candidates in resource-constrained regions. Thus, claims-based prognostic models represent a promising avenue for more accessible and large-scale prediction of AF ablation outcomes.

In this study, we propose to develop ML-based predictive models for outcomes of de novo AF ablation procedures using national-level claims data in the United States. Our goal is to evaluate an ML-derived risk prediction model for AF ablation patient outcomes. We hypothesize that ML models will be comparable to or exceed claim-based implementations of existing AF risk scores with respect to predictive power. Existing risk scores, including $CHADS_2$ and $CHA_2DS_2$-VASc, have achieved nontrivial improvements in predicting the outcomes of AF procedures (AUCs of 0.785 and 0.830, respectively, in a dataset consisting of 565 patients) [28]. Thus, in this study, we utilize $CHADS_2$ and $CHA_2DS_2$-VASc as a baseline to compare with our ML approaches. In addition, we also compare the performance between our ML models to a claims-based approximation of the CAAP-AF (modified CAAP-AF), a risk score specifically designed to predict AF ablation outcomes [12]. We also characterize outcomes by AF subtypes and sex and use different sets of parameters in the ML models to understand the contribution of individual factors to ML prediction performance.

## Methods

This research leveraged deidentified claims data sourced from the Merative MarketScan Research Medicare Databases (Merative, Inc.) between January 1, 2011, and December 31, 2021. MarketScan contains claims for individuals with Medicare Supplemental and Medicare Advantage plans.

### Patient Population

We analyzed Medicare claims data from January 1, 2011, to December 31, 2021, to identify patients who underwent AF ablation. Patients were included if they had a *Current Procedural Terminology* (*CPT*) code for AF ablation (93656) across either inpatient admission, inpatient services, and outpatient services tables in MarketScan. To ensure the accurate identification of AF ablation procedures, we required patients to have both *CPT* and a concurrent diagnosis of AF (*International Classification of Diseases, Ninth Revision* [*ICD-9*] code of "427.31" or *International Statistical Classification of Diseases, Tenth Revision* [*ICD-10*] code of "I48.X"). Each patient's medical history included all *ICD* codes from visits within 2 years before the initial occurrence of AF ablation within our dataset. While the 2-year lookback period captures baseline characteristics, claims data do not allow definitive confirmation that these represent truly de novo ablations, as patients may have undergone prior ablations before their enrollment period or outside the MarketScan database. Therefore, our cohort represents the best approximation of first-time AF ablation procedures available from administrative claims data.

We focused exclusively on Medicare beneficiaries for several reasons. First, the MarketScan database maintains separate patient identifiers for Medicare and commercial claims datasets, preventing integration of these patients. Second, the typical age for the first AF ablation is between 55 and 62 years [1,2,13], which is commonly covered by Medicare. Moreover, the substantial absence of postoperative outcomes for patients in the commercial database rendered it unsuitable for this study. The final cohort included 14,521 Medicare patients.

### Outcome and Subgroup Definitions

Our study's objective was to predict the binary outcome, success or failure, of AF ablation using patient demographics and prior medical history. Although the outpatient services table clearly documents the operation date for AF ablation, the inpatient admission and inpatient services tables only provide admission and discharge dates. To integrate the information across the 3 tables, we designated the admission date from the inpatient admission and inpatient services datasets as a surrogate for the AF ablation operation date in our analysis to maintain temporal coherence. Success was then defined as the absence of AF recurrence or repeat AF ablation between 6 and 12 months after the initial AF ablation procedure date, which is the standard interval before repeating an AF ablation according to current clinical practices [1,13]. To ensure the accurate identification of successful cases, we verified that all patients had at least 1 clinical follow-up visit within the first year after ablation.

The study also employed subgroup analysis by stratifying patients into 3 groups based on AF type: paroxysmal AF, persistent AF, and AF with flutter. This analysis was only possible after October 1, 2015, as *ICD-10* codes provide more detailed AF type distinctions compared to earlier *ICD-9* codes. We defined persistent AF as patients with *ICD-10* codes I48.1, I48.11, I48.19, I48.2, or I48.21. Note that this reflects current changes in the terminology of types of AF as it combines persistent AF and chronic AF. We defined paroxysmal AF as patients with *ICD-10* code of I48.0 or I48.20, and free of persistent AF. AF with atrial flutter were patients with any atrial flutter codes (*ICD-10*: I48.3 or I48.4).

### Data Processing

We constructed a comprehensive 2-year historical patient snapshot by linking records across the inpatient admission, inpatient services, and outpatient services tables using unique patient identifiers. For each patient, we extracted demographic variables (age, sex, region, and industry) at the time of the index ablation, along with the ablation date, failure date (if applicable), and all *ICD* codes within the 2 years preceding the index procedure. To standardize diagnostic codes across our study period, we used the *ICD-10* Lookup tool [30] to convert post-October 2015 *ICD-10* codes to their *ICD-9* equivalents. For computational efficiency and feature set manageability, we truncated all *ICD-9* codes to their first 3 digits, resulting in 785 *ICD* features and 19 demographic features. We used a binary measurement to denote whether or not a patient had the specific code within the 2 years prior to the initial ablation, thus avoiding extensive missing data.

We also calculated 2 established indices, the Charlson comorbidity index and the Elixhauser comorbidity index, to capture patients' comorbid conditions [31,32]. These indices used a weighted system based on specific conditions to provide a score, with higher values indicating more severe comorbidities.

For the subgroup analysis, we utilized three distinct datasets: (1) all the simplified 3-digit *ICD* codes, demographic information, and 2 established indices; (2) demographic data and 2 established indices; and (3) solely demographic information.

### Modeling

We used 2 popular supervised ML classifiers: logistic regression and extreme gradient boosting (XGBoost) [33]. Logistic regression computes the probability of a binary outcome by employing a logistic function (sigmoid curve) to transform the linear combination of input features into probabilities. This model is particularly advantageous due to its simplicity and interpretability, especially in scenarios where the relationship between input variables and the outcome is expected to be linear. To tune the logistic regression model, we implemented grid search with 5-fold stratified cross-validation AUC as the primary evaluation metric. We explored regularization strengths (C values) on a logarithmic scale (0.001, 0.01, 0.1, 1, 10, 100) to address potential overfitting concerns. Both L1 (Lasso) and L2 (Ridge) regularization penalties were investigated to determine the optimal feature selection. We evaluated multiple solvers ("liblinear," "lbfgs," "newton-cg," "sag," "saga") to identify the most computationally efficient optimization algorithm.

XGBoost represents a more sophisticated ML approach. XGBoost constructs multiple decision trees in a sequential manner, with each subsequent tree focusing on addressing the errors made by its predecessors. This method does not presuppose a linear relationship between input and output variables, offering greater flexibility and efficacy in dealing with larger and more intricate datasets. Despite its computational intensity, XGBoost is celebrated for its high efficiency and versatility, making it a potent tool in predictive modeling, especially in situations where the complexity of the data surpasses the capabilities of simpler models like logistic regression [27]. To tune the XGBoost hyperparameters, we implemented grid search with 5-fold stratified cross-validation with AUC as the primary evaluation metric. We explored a range of maximum depth values (3, 6, 9, 12, 15) to adequately capture complex feature interactions while avoiding overfitting. The learning rate varied across 0.01, 0.05, 0.1, and 0.2 to balance convergence speed and model accuracy, while the number of estimators was tested at 100, 200, 300, and 500 to determine the optimal number of boosting rounds.

The $CHADS_2$ and $CHA_2DS_2$-VASc risk scores have been widely used to predict stroke risk in patients with AF [10,11,28]. $CHADS_2$ is calculated using congestive heart failure, hypertension, age ≥75 years, diabetes, stroke (doubled), while $CHA_2DS_2$-VASc is computed using congestive heart failure, hypertension, age ≥75 (doubled) years, diabetes, stroke (doubled), vascular disease, age 65-74 years, and sex category (female). These risk scores more recently have been used to predict outcomes in patients with AF, heart failure, coronary artery disease, and postoperative AF undergoing cardiovascular surgical procedures [11,28]. $CHADS_2$ and $CHA_2DS_2$-VASc risk scores were shown to be useful predictors of adverse events after AF ablation [10]. In addition to $CHADS_2$ and $CHA_2DS_2$-VASc, we also evaluated a modified CAAP-AF, a risk score specifically designed to estimate the likelihood of remaining AF-free after ablation [12]. Due to the limitations of claims data, we only have information on coronary artery disease, age, AF type (persistent or longstanding, available only for patients after October 2015 using *ICD-10*), and sex. Left atrial diameter and the number of failed antiarrhythmic drugs are unavailable in MarketScan, which may impact our CAAP-AF comparison.

### Statistical Analysis

We compared patient characteristics between the groups using the Student *t* test for continuous variables and chi-square tests for categorical variables. Continuous variables were reported as mean (SD), while categorical variables were expressed as percentages.

We assessed the performance of the ML models and baseline risk scores using 5 metrics: AUC, area under the precision recall curve (AUPRC), precision, recall, and $F_1$-score (the harmonic mean of precision and recall). Optimal hyperparameters for the ML models were first identified through 5-fold cross-validation on the full dataset. To measure performance, we then employed bootstrap resampling with 500 iterations. In each iteration, the ML model was trained on a bootstrap sample (drawn with replacement from the full dataset) using these optimal hyperparameters and then evaluated on the out-of-bag observations (samples not included in that bootstrap sample). This procedure was used to generate performance distributions and 95% CIs. Statistical significance was assessed using 1-tailed paired *t* tests on the bootstrap distributions to test whether ML outperformed the clinical scores (H : XGBoost≤clinical score).

### Ethical Considerations

This study used commercially available data that have been deidentified. As such, the study was deemed exempt by Emory University Institutional Review Board.

## Results

We leveraged the Merative MarketScan Research Medicare Databases (Merative, Inc.) between January 1, 2011, and December 31, 2021. To allow for a 2-year medical history and 1-year outcome assessment, the analytic cohort included patients observed between January 1, 2013, and December 31, 2020.

The demographic and clinical profiles of the patients with AF are detailed in Tables 1 and 2. Our study cohort consisted of 14,521 patients, with an average age of 71.5 years (SD 5.31). Successful outcomes from AF ablation procedures were observed in 54.01% (n=7843) of the patients. Female patients constituted 39.94% (n=5800) of the study population. Clinically, 24.73% (n=3591) of the patients were diagnosed with concomitant atrial flutter. As shown in Table 2, the Elixhauser comorbidity index showed limited variance, with 92.89% (n=13,488) of the patients in the "≥2" category. The precise identification of patients with paroxysmal and persistent AF was limited, relative to the total cohort, due to the use of *ICD-9* instead of *ICD-10* prior to October 2015. A total of 7646 patients were identified using *ICD-10* codes for AF ablation, demonstrating a slightly reduced AF ablation success rate of 53.28% in comparison to the broader patient population. A subset of 6983 patients was categorized as having paroxysmal or persistent AF. Within this subset, 37.63% (n=2877) were diagnosed with paroxysmal AF, while 53.70% (n=4106) had persistent AF. The AF ablation success rates for paroxysmal and persistent AF were 52.55% and 53.90%, respectively.

**Table .** Baseline demographic characteristics of patients undergoing AF[a] ablation.[b]

| Demographic variable | Overall (N=14,521) | AF ablation success (n=7843) | AF ablation failure (n=6678) | *P* value |
|---|---|---|---|---|
| Age (y), mean (SD) | 71.5 (5.31) | 71.5 (5.27) | 71.6 (5.34) | .62 |
| Female, n (%) | 5800 (39.94) | 3118 (39.76) | 2682 (40.17) | .63 |
| Region, n (%) | | | | <.001 |
| Northeast | 2790 (19.21) | 1544 (19.69) | 1246 (18.66) | .12 |
| North Central | 4467 (30.76) | 2263 (28.85) | 2204 (33.00) | <.001 |
| South | 4733 (32.59) | 2599 (33.14) | 2134 (31.96) | .13 |
| West | 2393 (16.48) | 1360 (17.34) | 1033 (15.47) | .003 |
| Unknown | 138 (0.95) | 77 (0.98) | 61 (0.91) | .74 |
| Industry, n (%) | | | | <.001 |
| Oil and gas extraction, mining | 6 (0.04) | 5 (0.06) | 1 (0.01) | .30 |
| Manufacturing, non-durable goods | 3013 (20.75) | 1486 (18.94) | 1527 (22.87) | <.001 |
| Manufacturing, durable goods | 467 (3.21) | 254 (3.24) | 213 (3.19) | .91 |
| Transportation, communication, utilities | 1768 (12.18) | 1007 (12.84) | 761 (11.40) | .009 |
| Retail trade | 42 (0.29) | 22 (0.28) | 20 (0.30) | .95 |
| Finance, insurance, real estate | 661 (4.55) | 371 (4.73) | 290 (4.34) | .28 |
| Services | 2866 (19.74) | 1479 (18.86) | 1387 (20.77) | .004 |
| Agriculture, forestry, fishing | 4 (0.03) | 2 (0.03) | 2 (0.03) | >.99 |
| Construction | 33 (0.23) | 20 (0.26) | 13 (0.19) | .56 |
| Wholesale | 54 (0.37) | 37 (0.47) | 17 (0.25) | .05 |
| Unknown | 5607 (38.61) | 3160 (40.29) | 2447 (36.64) | <.001 |

[a]AF: atrial fibrillation.

[b]Industry is categorized based on the employer responsible for the claim payment, and regions follow the Census Bureau's regional definitions.

**Table .** Baseline clinical characteristics of patients in sample undergoing AF[a] ablation.[b]

| Clinical variable | Overall (N=14,521), n (%) | AF ablation success (n=7843), n (%) | AF ablation failure (n=6678), n (%) | *P* value |
|---|---|---|---|---|
| Charlson comorbidity index | | | | .29 |
| 0 | 4371 (30.10) | 2375 (30.28) | 1996 (29.89) | .61 |
| 1 | 4295 (29.58) | 2277 (29.03) | 2018 (30.22) | .87 |
| ≥2 | 5855 (40.32) | 3191 (40.68) | 2664 (39.89) | .61 |
| Elixhauser comorbidity index | | | | .45 |
| 0 | 44 (0.30) | 24 (0.31) | 20 (0.30) | >.99 |
| 1 | 989 (6.81) | 515 (6.57) | 474 (7.10) | >.99 |
| ≥2 | 13,488 (92.89) | 7304 (93.13) | 6184 (92.60) | >.99 |
| Both atrial flutter and AF (*ICD-9*[c] and *ICD-10*[d]) | 3591 (24.73) | 1963 (25.03) | 1628 (24.38) | .38 |
| Patients with *ICD-10* | 7646 (52.65) | 4074 (51.94) | 3572 (53.49) | .51 |
| Paroxysmal AF (*ICD-10* only) | 2877 (37.63) | 1512 (37.11) | 1365 (38.21) | .33 |
| Persistent AF (*ICD-10* only) | 4106 (53.70) | 2213 (54.32) | 1893 (53.00) | .26 |
| Unspecified AF | 663 (8.67) | 349 (8.57) | 314 (8.79) | .76 |

[a]AF: atrial fibrillation.

[b]The paroxysmal and persistent AF only exists in the *ICD-10* space, of which the overall *ICD-10* population is 7646, success population is 4074, and failure population is 3572.

[c]*ICD-9*: *International Classification of Diseases, Ninth Revision*.

[d]*ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*.

Table 3 shows the comparative performance of XGBoost, $CHADS_2$, and $CHA_2DS_2$ -VASc of our entire study cohort. XGBoost consistently outperformed logistic regression in all analyses; therefore, only XGBoost results are presented for brevity. The full comparison between XGBoost and logistic regression is available in Multimedia Appendix 1. The XGBoost model exhibited modest predictive capability with an AUC of 0.528 for the overall population. It performed slightly better in male (AUC=0.529) than in female patients (AUC=0.521). The model achieved balanced performance with an $F_1$-score of 0.581 and recall of 0.608, indicating that it captures most positive cases while maintaining reasonable precision at 0.556. Male patients showed slightly higher recall (0.614) than female patients (0.600).

**Table .** Performance comparison between XGBoost[a] and $CHADS_2$[b] and $CHA2_2DS2_2$-VASc[c] risk scores stratified by sex.

| Metric | XGBoost[d] | $CHADS_2$ | $CHA_2DS_2$-VASc |
|---|---|---|---|
| Population (n=14,521) | | | |
| AUC[e] | 0.528[f] (0.519 - 0.533) | 0.498 | 0.498 |
| AUPRC[g] | 0.562[f] (0.545 - 0.578) | 0.536 | 0.539 |
| $F_1$-score | 0.581 (0.569 - 0.594) | 0.436 | 0.644 |
| Precision | 0.556[f] (0.542 - 0.570) | 0.533 | 0.540 |
| Recall | 0.608 (0.585 - 0.632) | 0.368 | 0.799 |
| Female (n=5800) | | | |
| AUC | 0.521[f] (0.510 - 0.532) | 0.498 | 0.500 |
| AUPRC | 0.558[f] (0.533 - 0.582) | 0.536 | 0.538 |
| $F_1$-score | 0.575 (0.556 - 0.593) | 0.436 | 0.698 |
| Precision | 0.552[f] (0.530 - 0.574) | 0.533 | 0.538 |
| Recall | 0.600 (0.568 - 0.632) | 0.368 | 0.995 |
| Male (n=8721) | | | |
| AUC | 0.529[f] (0.520 - 0.539) | 0.498 | 0.498 |
| AUPRC | 0.566[f] (0.546 - 0.588) | 0.541 | 0.541 |
| $F_1$-score | 0.585 (0.568 - 0.601) | 0.410 | 0.599 |
| Precision | 0.559[f] (0.540 - 0.578) | 0.538 | 0.542 |
| Recall | 0.614 (0.582 - 0.644) | 0.331 | 0.669 |

[a]XGBoost: extreme gradient boosting.

[b]$CHADS_2$: congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, and stroke.

[c]$CHA_2DS_2$-VASc: congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age, and sex category.

[d]Cell values for XGBoost report average and the 95% CI in parentheses.

[e]AUC: area under the receiver operating characteristic curve.

[f]*P*<.001 (XGBoost vs both clinical scores).

[g]AUPRC: area under the precision recall curve.

Despite its moderate predictive power, the XGBoost model consistently outperformed both $CHADS_2$ and $CHA_2DS_2$ -VASc scores across all patient cohorts in terms of AUC and AUPRC. Both risk scores ($CHADS_2$ and $CHA_2$ $DS_2$ -VASc) performed poorly with AUC values worse than random chance (<0.5) except for $CHA_2DS_2$-VASc in the female subgroup (AUC=0.5). $CHADS_2$ had poor recall (0.368) and low $F_1$-scores (0.436), missing most positive cases. While $CHA_2DS_2$-VASc demonstrated high recall (0.799), particularly in female patients (0.995), the lower precision of 0.540 and AUC below 0.5 suggest that the score's high sensitivity produces a higher false positive rate.

Table 4 presents a comparative analysis of the XGBoost and the modified CAAP-AF risk scores for the *ICD-10* cohort. XGBoost outperformed the modified CAAP-AF risk score with an overall AUC of 0.544 and AUPRC of 0.567 and consistent subgroup performance (female patients: AUC 0.543, AUPRC 0.569; male patients: AUC 0.545, AUPRC 0.567). In contrast, the modified CAAP-AF risk score achieved an overall AUC of 0.505, rising slightly to 0.511 in male patients but performing no better than random chance in female patients. While the modified CAAP-AF risk score exhibited high recall (0.999), capturing nearly all positive cases, it came at the cost of lower precision (0.533). XGBoost achieved better precision (0.552), albeit at a lower recall (0.793) and $F_1$-score. All differences between the models were statistically significant.

**Table .** Performance comparison between XGBoost[a] (ML[b] model) and modified CAAP-AF[c] risk score stratified by sex.

| Metric | XGBoost[d] | Modified CAAP-AF |
|---|---|---|
| *ICD-10*[e] population (n=7646) | | |
| AUC[f] | 0.544[g] (0.535 - 0.553) | 0.505 |
| AUPRC[h] | 0.567[g] (0.545 - 0.590) | 0.537 |
| $F_1$-score | 0.651 (0.627 - 0.672) | 0.695[g] |
| Precision | 0.552[g] (0.535 - 0.572) | 0.533 |
| Recall | 0.793 (0.709 - 0.867) | 0.999[g] |
| Female (n=3161) | | |
| AUC | 0.543[g] (0.516 - 0.573) | 0.500 |
| AUPRC | 0.569[g] (0.533 - 0.605) | 0.533 |
| $F_1$-score | 0.645 (0.615 - 0.672) | 0.694[g] |
| Precision | 0.550[g] (0.519 - 0.578) | 0.531 |
| Recall | 0.783 (0.701 - 0.864) | 1.000[g] |
| Male (n=4485) | | |
| AUC | 0.545[g] (0.522 - 0.569) | 0.511 |
| AUPRC | 0.567[g] (0.540 - 0.596) | 0.542 |
| $F_1$-score | 0.655 (0.625 - 0.678) | 0.696[g] |
| Precision | 0.554[g] (0.533 - 0.579) | 0.535 |
| Recall | 0.801 (0.717 - 0.875) | 0.999[g] |

[a]XGBoost: extreme gradient boosting.

[b]ML: machine learning.

[c]CAAP-AF: coronary artery disease, left atrial diameter, age, AF, antiarrhythmic drugs, and female sex category.

[d]Cell values for XGBoost report average and the 95% CI in parentheses.

[e]*ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*.

[f]AUC: area under the receiver operating characteristic curve.

[g]$P$<.001 for comparison between XGBoost and modified CAAP-AF.

[h]AUPRC: area under the precision recall curve.

Table 5 presents the predictive model performance across atrial arrhythmia subgroups: paroxysmal AF, persistent AF, and AF with atrial flutter. A total of 3 feature sets were compared: *ICD* codes with demographics and comorbidity indices, demographics and comorbidity indices, and demographics only. On the entire *ICD-10* population, incorporating all the features (*ICD* codes with demographics and comorbidity indices) achieved the best performance across all 5 metrics when compared to the other 2 feature sets, with AUC of 0.544, AUPRC of 0.567, $F_1$-score of 0.652, precision of 0.551, and recall of 0.798.

**Table .** Predictive performance by clinical and demographic predictors across atrial arrhythmia subgroups.[a]

| Metric | *ICD*[b]+demographic+comorbidity indices, average (95% CI) | Demographic+comorbidity indices, average (95% CI) | Demographic only, average (95% CI) |
|---|---|---|---|
| Paroxysmal AF[c] (n=2877) | | | |
| AUC[d] | 0.538 (0.523 - 0.553) | 0.520 (0.530 - 0.558) | 0.532 (0.517 - 0.547) |
| AUPRC[e] | 0.557 (0.520 - 0.596) | 0.564 (0.529 - 0.595) | 0.547 (0.515 - 0.582) |
| $F_1$-score | 0.563 (0.531 - 0.593) | 0.596 (0.540 - 0.639) | 0.645 (0.581 - 0.680) |
| Precision | 0.551 (0.514 - 0.585) | 0.548 (0.513 - 0.582) | 0.541 (0.509 - 0.570) |
| Recall | 0.576 (0.520 - 0.629) | 0.660 (0.525 - 0.789) | 0.808 (0.620 - 0.948) |
| Persistent AF (n=4106) | | | |
| AUC | 0.525 (0.512 - 0.537) | 0.518 (0.504 - 0.531) | 0.524 (0.510 - 0.537) |
| AUPRC | 0.561 (0.532 - 0.592) | 0.552 (0.522 - 0.582) | 0.557 (0.529 - 0.586) |
| $F_1$-score | 0.575 (0.550 - 0.596) | 0.626 (0.586 - 0.659) | 0.658 (0.612 - 0.689) |
| Precision | 0.554 (0.525 - 0.582) | 0.549 (0.522 - 0.574) | 0.545 (0.520 - 0.573) |
| Recall | 0.598 (0.553 - 0.641) | 0.731 (0.622 - 0.821) | 0.834 (0.677 - 0.947) |
| *ICD-10*[f], with AF (n=1503) | | | |
| AUC | 0.528 (0.506 - 0.549) | 0.514 (0.493 - 0.535) | 0.517 (0.497 - 0.537) |
| AUPRC | 0.564 (0.516 - 0.609) | 0.555 (0.512 - 0.602) | 0.558 (0.513 - 0.605) |
| $F_1$-score | 0.600 (0.551 - 0.644) | 0.607 (0.552 - 0.655) | 0.611 (0.556 - 0.655) |
| Precision | 0.556 (0.508 - 0.601) | 0.547 (0.507 - 0.589) | 0.693 (0.548 - 0.823) |
| Recall | 0.657 (0.551 - 0.770) | 0.688 (0.552 - 0.815) | 0.550 (0.509 - 0.597) |
| *ICD-10* population (n=7646) | | | |
| AUC | 0.544 (0.535 - 0.553) | 0.533 (0.523 - 0.542) | 0.536 (0.528 - 0.545) |
| AUPRC | 0.567 (0.545 - 0.589) | 0.556 (0.532 - 0.579) | 0.559 (0.536 - 0.581) |
| $F_1$-score | 0.652 (0.625 - 0.672) | 0.621 (0.595 - 0.645) | 0.645 (0.610 - 0.672) |
| Precision | 0.551 (0.531 - 0.573) | 0.550 (0.530 - 0.570) | 0.548 (0.528 - 0.570) |
| Recall | 0.798 (0.713 - 0.871) | 0.714 (0.644 - 0.796) | 0.787 (0.681 - 0.878) |

[a]This population only includes patients who had their first atrial fibrillation ablation in or after October 2015. Predictors include *ICD* codes of patients' past medical history and demographic variables (region, sex, age, and industry).

[b]*ICD*: *International Classification of Diseases*.

[c]AF: atrial fibrillation.

[d]AUC: area under the receiver operating characteristic curve.

[e]AUPRC: area under the precision recall curve.

[f]*ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*.

Within the atrial arrhythmia subgroups, models incorporating all features consistently achieved the highest AUC and AUPRC across all 3 subgroups. However, performance patterns for $F_1$-score and recall varied by subgroup. For paroxysmal AF and persistent AF, the full model also achieved the highest precision (0.551 and 0.554, respectively), but the models with demographics only had better recall (0.808 and 0.834, respectively) and $F_1$-scores (0.645 and 0.658, respectively). For AF with atrial flutter, the model with demographics only achieved the highest $F_1$-score (0.611) and precision (0.693), whereas the model with demographics and comorbidity indices achieved the highest recall (0.688).

## Discussion

### Principal Findings

In this study, we developed ML models that predict the outcomes of de novo AF ablation procedures. Our XGBoost model demonstrated statistically significant improved performance compared to 3 different claim-based implementations of clinical risk scores ($CHADS_2$, $CHA_2DS_2$-VASc, and a limited, modified CAAP-AF without left atrial diameter and the number of failed antiarrhythmic drugs) in all patient and sex subgroups in terms of AUC and AUPRC. While the 2 risk scores achieved higher recall than

XGBoost, they demonstrated lower precision and weaker discrimination (near random chance). However, XGBoost's predictive ability for outcomes after AF ablation was found to be lower in female patients than it was in male patients or in the entire population. There was no difference in AUC when comparing $CHADS_2$ to $CHA_2DS_2$-VASc risk scores for outcomes after AF ablation except for female patients, where $CHA_2DS_2$-VASc performs better than $CHADS_2$.

When comparing outcomes across different AF subtypes (paroxysmal, persistent, or AF with atrial flutter), we observed heterogeneous patterns in the value of adding *ICD* code features. For persistent AF and AF with atrial flutter, the models incorporating *ICD* code features demonstrated superior discriminative power (AUC and AUPRC) compared to models using either demographic or clinical variables alone or those combined with comorbidity indices (Charlson comorbidity index and the Elixhauser comorbidity index). However, in the paroxysmal AF subgroup, the model using only demographics and comorbidity indices slightly outperformed the full model with *ICD* codes in terms of AUPRC but not AUC. Additionally, models using demographics only consistently achieved higher recall across all subgroups at the expense of lower precision and overall discriminative performance (AUC and AUPRC), revealing a trade-off between sensitivity and specificity in feature selection. The use of these ML models may be useful in clinical practice in patient selection for AF ablation in the future.

## Comparison to Prior Work

Claims data present challenges for outcome prediction, despite being readily available. Previous clinical models for predicting AF ablation success have reported an AUC ranging from 0.55 to 0.65, with only 3 models achieving an AUC of 0.75 [4,5,12]. In other studies, $CHADS_2$ and $CHA_2DS_2$-VASc achieved an AUC of 0.785 and 0.830, respectively, in predicting complications after AF ablation [6]. However, in our study, $CHADS_2$ and $CHA_2DS_2$-VASc only achieved an AUC of 0.498 - 0.5, performing almost worse than random guessing. It is important to note that while $CHADS_2$ and $CHA_2DS_2$-VASc have been used for predicting procedural outcomes [4-6], they were originally designed to estimate stroke risk rather than ablation recurrence, and thus their lower performance in this study potentially reflects use outside of the intended purpose rather than a failure of the scores themselves.

The modified CAAP-AF reached AUC greater than 0.650 [12] with the data from its original study, yet in our implementation, it achieved no better than 0.511. However, the CAAP-AF score used in our study was a modified, claims-based approximation that excluded left atrial diameter and the number of failed antiarrhythmic drugs, as these are not available in claims data. Therefore, our comparison does not represent a true head-to-head evaluation of the original CAAP-AF model, and the ML model's advantage should be interpreted with this limitation in mind.

These findings highlight the significant difficulty in predicting AF ablation success and failure using claims data, reflecting broader challenges in health care outcome prediction where administrative databases consistently underperform compared to clinical models due to the absence of key physiological and procedural variables, a pattern observed across multiple medical specialties and intervention types [34,35]. In contrast, our ML models achieved AUCs of 0.521 - 0.529, showing marginal improvement.

Despite the modest predictive performance of the ML models, our claims-based approach has significant potential for standardization across health care systems, as it relies on widely used *ICD* and *CPT* coding systems rather than institution-specific EHR implementations. However, adoption barriers remain, including variations in coding practices across institutions, the challenge of integrating predictive tools into clinical workflows, and potential resistance from clinicians who may prioritize clinical judgment over algorithmic recommendations. Given the relatively low AUC values observed, these models should be viewed as a foundational step toward using claims data to predict the outcomes of AF ablation procedures, rather than as tools ready for clinical deployment.

Beyond demonstrating that ML models outperform traditional risk scores, we conducted an analysis to understand what types of features should be included in the ML models across clinically relevant AF subgroups. We evaluated three feature sets: (1) demographic information alone; (2) demographics plus comorbidity indices; and (3) the full features incorporating *ICD* codes, demographics, and comorbidity indices. These were tested across 3 clinically distinct subgroups (paroxysmal AF, persistent AF, and AF with atrial flutter) identifiable only through *ICD-10* coding, yielding 16 unique ML models. Across persistent AF and AF with atrial flutter subgroups, ML models performed best when including *ICD* codes as features, highlighting the importance of diagnostic coding data. Among the 3 subgroups (paroxysmal AF, persistent AF, and patients with atrial flutter), the ML models performed best for patients with paroxysmal AF, and patients with persistent AF had the least success. The entire *ICD-10* population achieved the highest overall AUC compared to other subgroups, which was likely due to the larger sample size.

## Future Directions

Our findings demonstrate that ML models using *ICD* codes to estimate AF ablation procedural outcomes are robust and valid across populations. However, the model's current predictive power in this study remains insufficient for clinical decision-making. Improvement of outcome predictions for AF ablation using ML has the potential for widespread use in research and clinical practice to determine optimal patient selection for AF ablation and the management of patients with AF. Advances in artificial intelligence and ML technology have an ability to rapidly analyze and synthesize innumerable variables to predict outcomes of AF ablation and discover new patterns of clinical variables that greatly surpass prior conventional methods of gaging success. These findings will be important to consider, as health care policymakers struggle to allocate limited resources to as many patients as possible and search for ways to improve patient outcomes. ML technologies will play increasingly more important roles in medicine with future advances as we better learn how to incorporate ML for

better health care resource allocation as well as improvements in clinical practice and patient outcomes.

Several specific clinical implementation scenarios could leverage these predictive tools to enhance AF ablation care delivery. An important deployment consideration is the metric to optimize, as our findings revealed a trade-off between precision and recall. For population health monitoring, quality improvement initiatives, or patient counseling, high-recall models may be preferred. Conversely, for resource allocation decisions such as prioritizing ablation slots during periods of limited procedure capacity, high-precision models would be more appropriate to minimize false positives. Clinicians could use model predictions to provide patients with more personalized success probability estimates during shared decision-making discussions, helping patients make more informed treatment choices. Alternatively, these models could guide the development of alternative treatment pathways or enhanced monitoring protocols for patients with consistently lower predicted success rates. Future research should focus on developing implementation frameworks that appropriately balance algorithmic predictions with clinical judgment and metric selection based on clinical context while ensuring equitable access to AF ablation across diverse patient populations.

## Limitations

First, our study relied exclusively on Medicare Advantage and Medicare Supplemental claims, which skews the cohort toward older patients. Although first ablations often occur between ages 55 - 62 years, our findings may not be generalizable to younger populations with different comorbidity profiles and procedural outcomes. The etiology and pathophysiology of AF may differ between younger and older patients, which could affect both the predictive variables and outcomes in ways that our models may not capture. Future work should validate and potentially recalibrate these models in younger and more diverse populations to ensure broader clinical utility.

Second, as with all administrative data, coding errors and inconsistencies are possible. We mitigated this by truncating *ICD* codes into broader categories, incorporating 2 established comorbidity indices (Charlson comorbidity index and the Elixhauser comorbidity index), and requiring that all patients had a documented AF diagnosis before ablation. Despite these steps, misclassification could still reduce model performance. Moreover, truncating *ICD-9* codes to the first 3 digits may also have reduced diagnostic specificity. This limitation may explain our unexpected finding that the model using only demographics and comorbidity indices slightly outperformed the full model with *ICD* codes in the paroxysmal AF subgroup. The truncated *ICD* codes may have introduced noise rather than signal for this subgroup, particularly if patients with paroxysmal AF have less diverse billing code profiles making the additional *ICD* code features less informative. Future analysis may mitigate the issue by integrating claims with richer data sources to cross-validate the information.

Third, while we aimed to study de novo AF ablation procedures, administrative claims data have inherent limitations in both identifying first-time ablations and measuring their outcomes. Although we identified the initial occurrence of AF ablation within our dataset, we cannot definitively exclude patients who may have undergone prior ablations before their enrollment in the database or at facilities not captured in MarketScan. Furthermore, our outcome definition may be subject to misclassification as we are using billing codes as a proxy for clinical recurrence. Asymptomatic or unrecorded recurrences could be missed (falsely classified as success), while unrelated visits coded with previous AF could be incorrectly classified as failures. Additionally, patients with undetected prior ablations may have different recurrence trajectories than true first-time procedures, further complicating outcome assessment. AF recurrence is best confirmed with secondary data sources such as Holter monitoring or electrocardiogram data.

Fourth, claims data lack important clinical variables known to influence AF ablation outcomes, such as left atrial size, ejection fraction, specific antiarrhythmic medications, and procedural details (catheter type, ablation strategy). This limitation likely contributed to our models' modest predictive performance compared to clinical prediction models that incorporate these variables. Additionally, the limited variance in the Elixhauser comorbidity index, where 92.89% (n=13,488) of patients fell into a single category (≥2), reduced its discriminative value and may explain why adding comorbidity indices to demographic variables resulted in minimal or slightly negative effects on model performance in some subgroups. While we cannot address this limitation within our study design, future research could explore hybrid approaches that combine claims data with targeted clinical data collection for key predictive variables. However, we note this may limit the scalability and standardization advantages that motivated our claims-based approach.

Finally, given the proprietary nature of MarketScan data, direct replication is constrained. To enhance transparency and reproducibility, we documented our data source, inclusion and exclusion criteria, billing codes, and potential confounders and released the analytic code in a public GitHub repository to facilitate replication [36]. This enables researchers with access to similar claims databases to replicate our methodology, though exact replication would require the same data source.

## Conclusions

In this study, we developed and evaluated ML models using MarketScan claims data to predict 1-year AF ablation outcomes. Across the overall cohort and sex-stratified groups, ML models modestly but consistently outperformed claim-based implementations of established clinical risk scores. In the *ICD-10* subset, incorporating *ICD* diagnostic codes improved performance relative to the models using only demographic and comorbidity indices over most subgroups. Our findings demonstrate the limitations of ML approaches when applied to claims data that lack key clinical variables, such as left atrial size, ejection fraction, and medication details. The modest predictive performance indicates that current claims-based models are insufficient for individual clinical decision-making. Despite these constraints, our work establishes that standardized, population-level outcome prediction using nationally available administrative data is technically feasible, providing capability

that could complement existing clinical tools for health system quality monitoring and research applications. These results contribute important insights into the potential and limitations of claims-based prediction models for population-level analyses and comparative effectiveness research.

## Acknowledgments

No generative artificial intelligence (GenAI) models (including large language models such as ChatGPT) were used in the conduct of this research, data analysis, or original writing of this manuscript. Code auto-completion tools were used during programming, with all generated code manually reviewed. Limited GenAI assistance was used to refine language during manuscript revision. We used the Python scikit-learn module to perform logistic regression and extreme gradient boosting, which formed the body of this research [37].

## Funding

This work was supported by the National Institutes of Health (grant NHLBI #R21HL156184; principal investigator: VSH).

## Data Availability

The Merative MarketScan Research Medicare database is proprietary and not publicly accessible. Interested researchers may obtain access directly from Merative under a licensing agreement.

## Authors' Contributions

Conceptualization: JCH, MSL, VSH, YL
Data curation: MOO, VSH, YL
Formal analysis: JCH, KAW, MOO, MSL, VSH, YL
Funding acquisition: JCH, KAW, VSH
Investigation: YL
Methodology: JCH, MOO, VSH, YL
Project administration: VSH
Supervision: JCH, KAW, VSH
Writing – original draft: YL
Writing – review & editing: JCH, KAW, MOO, MSL, VSH

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Logistic regression results.
[DOCX File, 3252 KB - cardio_v9i1e77380_app1.docx ]

## References

1. Tzeis S, Gerstenfeld EP, Kalman J, et al. 2024 European Heart Rhythm Association/Heart Rhythm Society/Asia Pacific Heart Rhythm Society/Latin American Heart Rhythm Society expert consensus statement on catheter and surgical ablation of atrial fibrillation. Europace 2024 Mar 30;26(4):euae043. [doi: 10.1093/europace/euae043] [Medline: 38587017]
2. Martin SS, Aday AW, Almarzooq ZI, et al. 2024 heart disease and stroke statistics: a report of US and global data from the American Heart Association. Circulation 2024 Feb 20;149(8):e347-e913. [doi: 10.1161/CIR.0000000000001209] [Medline: 38264914]
3. Andrade JG, Champagne J, Dubuc M, et al. Cryoballoon or radiofrequency ablation for atrial fibrillation assessed by continuous monitoring: a randomized clinical trial. Circulation 2019 Nov 26;140(22):1779-1788. [doi: 10.1161/CIRCULATIONAHA.119.042622] [Medline: 31630538]
4. Zheng JY, Li DT, Qıu YG, et al. CHA2DS2-VASc score as a prognostic indicator in patients with atrial fibrillation undergoing coronary stenting. Turk J Med Sci 2022 Aug;52(4):1103-1110. [doi: 10.55730/1300-0144.5413] [Medline: 36326398]
5. Mulder MJ, Kemme MJB, Hopman L, et al. Comparison of the predictive value of ten risk scores for outcomes of atrial fibrillation patients undergoing radiofrequency pulmonary vein isolation. Int J Cardiol 2021 Dec 1;344:103-110. [doi: 10.1016/j.ijcard.2021.09.029] [Medline: 34555444]
6. Jacobs V, May HT, Bair TL, et al. The impact of risk score (CHADS2 versus CHA2DS2-VASc) on long-term outcomes after atrial fibrillation ablation. Heart Rhythm 2015 Apr;12(4):681-686. [doi: 10.1016/j.hrthm.2014.12.034] [Medline: 25546809]

7. Kornej J, Hindricks G, Arya A, Sommer P, Husser D, Bollmann A. The APPLE score—a novel score for the prediction of rhythm outcomes after repeat catheter ablation of atrial fibrillation. PLoS One 2017;12(1):e0169933. [doi: 10.1371/journal.pone.0169933] [Medline: 28085921]
8. Kosich F, Schumacher K, Potpara T, Lip GY, Hindricks G, Kornej J. Clinical scores used for the prediction of negative events in patients undergoing catheter ablation for atrial fibrillation. Clin Cardiol 2019 Feb;42(2):320-329. [doi: 10.1002/clc.23139] [Medline: 30578568]
9. Shuvy M, Zwas DR, Keren A, Gotsman I. Value of the CHA2 DS2 -VASc score for predicting outcome in patients with heart failure. ESC Heart Fail 2020;7(5):2553-2560. [doi: 10.1002/ehf2.12831] [Medline: 32614479]
10. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the Euro Heart Survey on atrial fibrillation. Chest 2010 Feb;137(2):263-272. [doi: 10.1378/chest.09-1584] [Medline: 19762550]
11. Chen JY, Zhang AD, Lu HY, Guo J, Wang FF, Li ZC. CHADS2 versus CHA2DS2-VASc score in assessing the stroke and thromboembolism risk stratification in patients with atrial fibrillation: a systematic review and meta-analysis. J Geriatr Cardiol 2013 Sep;10(3):258-266. [doi: 10.3969/j.issn.1671-5411.2013.03.004] [Medline: 24133514]
12. Winkle RA, Jarman JWE, Mead RH, et al. Predicting atrial fibrillation ablation outcome: the CAAP-AF score. Heart Rhythm 2016 Nov;13(11):2119-2125. [doi: 10.1016/j.hrthm.2016.07.018] [Medline: 27435586]
13. Van Gelder IC, Rienstra M, Bunting KV, et al. 2024 ESC Guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2024 Sep 29;45(36):3314-3414. [doi: 10.1093/eurheartj/ehae176] [Medline: 39210723]
14. Clarnette JA, Brooks AG, Mahajan R, et al. Outcomes of persistent and long-standing persistent atrial fibrillation ablation: a systematic review and meta-analysis. Europace 2018 Nov 1;20(FI_3):f366-f376. [doi: 10.1093/europace/eux297] [Medline: 29267853]
15. Ngo L, Ali A, Ganesan A, Woodman R, Adams R, Ranasinghe I. Ten-year trends in mortality and complications following catheter ablation of atrial fibrillation. Eur Heart J Qual Care Clin Outcomes 2022 Jun 6;8(4):398-408. [doi: 10.1093/ehjqcco/qcab102] [Medline: 34982824]
16. Ngo L, Ali A, Ganesan A, Woodman R, Adams R, Ranasinghe I. Gender differences in complications following catheter ablation of atrial fibrillation. Eur Heart J Qual Care Clin Outcomes 2021 Sep 16;7(5):458-467. [doi: 10.1093/ehjqcco/qcab035] [Medline: 33963402]
17. Abdur Rehman K, Wazni OM, Barakat AF, et al. Life-threatening complications of atrial fibrillation ablation: 16-year experience in a large prospective tertiary care cohort. JACC Clin Electrophysiol 2019 Mar;5(3):284-291. [doi: 10.1016/j.jacep.2018.11.013] [Medline: 30898229]
18. Feeny AK, Chung MK, Madabhushi A, et al. Artificial intelligence and machine learning in arrhythmias and cardiac electrophysiology. Circ Arrhythm Electrophysiol 2020 Aug;13(8):e007952. [doi: 10.1161/CIRCEP.119.007952] [Medline: 32628863]
19. Nygårds ME, Hulting J. An automated system for ECG monitoring. Comput Biomed Res 1979 Apr;12(2):181-202. [doi: 10.1016/0010-4809(79)90015-6] [Medline: 371910]
20. Dey D, Slomka PJ, Leeson P, et al. Artificial intelligence in cardiovascular imaging. J Am Coll Cardiol 2019 Mar;73(11):1317-1335. [doi: 10.1016/j.jacc.2018.12.054]
21. Huecker M, Schutzman C, French J, et al. Accurate modeling of ejection fraction and stroke volume with mobile phone auscultation: prospective case-control study. JMIR Cardio 2024 Jun 26;8:e57111. [doi: 10.2196/57111] [Medline: 38924781]
22. Müller S, Diekmann S, Wenzel M, et al. Combining machine learning with real-world data to identify gaps in clinical practice guidelines: feasibility study using the prospective German Stroke Registry and the national acute ischemic stroke guidelines. JMIR Med Inform 2025 Jul 11;13:e69282. [doi: 10.2196/69282] [Medline: 40653745]
23. Alhumaidi NH, Dermawan D, Kamaruzaman HF, Alotaiq N. The use of machine learning for analyzing real-world data in disease prediction and management: systematic review. JMIR Med Inform 2025 Jun 19;13:e68898. [doi: 10.2196/68898] [Medline: 40537090]
24. MacKay EJ, Stubna MD, Chivers C, et al. Application of machine learning approaches to administrative claims data to predict clinical outcomes in medical and surgical patient populations. PLoS One 2021;16(6):e0252585. [doi: 10.1371/journal.pone.0252585] [Medline: 34081720]
25. Deng H, Shantsila A, Xue Y, et al. Using the MB-LATER score for predicting arrhythmia outcome after catheter ablation for atrial fibrillation: the Guangzhou atrial fibrillation project. Int J Clin Pract 2018 Nov;72(11):e13247. [doi: 10.1111/ijcp.13247] [Medline: 30144238]
26. Yuan N, Duffy G, Dhruva SS, et al. Deep learning of electrocardiograms in sinus rhythm from US veterans to predict atrial fibrillation. JAMA Cardiol 2023 Dec 1;8(12):1131-1139. [doi: 10.1001/jamacardio.2023.3701] [Medline: 37851434]
27. Tseng AS, Noseworthy PA. Prediction of atrial fibrillation using machine learning: a review. Front Physiol 2021;12:752317. [doi: 10.3389/fphys.2021.752317] [Medline: 34777014]
28. Chao TF, Lin YJ, Tsao HM, et al. CHADS(2) and CHA(2)DS(2)-VASc scores in the prediction of clinical outcomes in patients with atrial fibrillation after catheter ablation. J Am Coll Cardiol 2011 Nov 29;58(23):2380-2385. [doi: 10.1016/j.jacc.2011.08.045] [Medline: 22115643]

29. Levinson RT, Paul C, Meid AD, Schultz JH, Wild B. Identifying predictors of heart failure readmission in patients from a statutory health insurance database: retrospective machine learning study. JMIR Cardio 2024 Jul 23;8:e54994. [doi: 10.2196/54994] [Medline: 39042456]
30. ICD code lookup website. Lookup IC. 2024. URL: https://icdcodelookup.com [accessed 2025-12-16]
31. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
32. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care 1998 Jan;36(1):8-27. [doi: 10.1097/00005650-199801000-00004] [Medline: 9431328]
33. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Presented at: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13-17, 2016; San Francisco, CA, USA p. 785-794. [doi: 10.1145/2939672.2939785]
34. Desai RJ, Wang SV, Vaduganathan M, Evers T, Schneeweiss S. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. JAMA Netw Open 2020 Jan 3;3(1):e1918962. [doi: 10.1001/jamanetworkopen.2019.18962] [Medline: 31922560]
35. Zeltzer D, Balicer RD, Shir T, Flaks-Manov N, Einav L, Shadmi E. Prediction accuracy with electronic medical records versus administrative claims. Med Care 2019 Jul;57(7):551-559. [doi: 10.1097/MLR.0000000000001135] [Medline: 31135691]
36. yijunliu23/AFA-Claims-CodeRelease. GitHub. URL: https://github.com/yijunliu23/AFA-Claims-CodeRelease [accessed 2025-12-23]
37. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830. [doi: 10.5555/1953048.2078195]


## Abbreviations

**AF:** atrial fibrillation
**AUC :** area under the receiver operating characteristic curve
**AUPRC:** area under the precision recall curve
**CAAP-AF:** coronary artery disease, left atrial diameter, age, AF, antiarrhythmic drugs, and female sex category
**$CHA_2DS_2$-VASc:** congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age, and sex category
**$CHADS_2$:** congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, and stroke
***CPT*:** *Current Procedural Terminology*
**EHR:** electronic health record
***ICD*:** *International Classification of Disease*
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
***ICD-9*:** *International Classification of Diseases, Ninth Revision*
**ML:** machine learning
**XGBoost:** extreme gradient boosting


*Edited by A Coristine; submitted 12.05.25; peer-reviewed by C Gissel, VG Yogeshappa; revised version received 04.12.25; accepted 04.12.25; published 31.12.25.*

*Please cite as:*
*Liu Y, Oloko-Oba M, Wood KA, Lloyd MS, Ho JC, Hertzberg VS*
*Predicting Atrial Fibrillation Ablation Outcomes: Machine Learning Model Development and Validation Using a Large Administrative Claims Database*
*JMIR Cardio 2025;9:e77380*
*URL: https://cardio.jmir.org/2025/1/e77380*
*doi:10.2196/77380*

# Conversational AI Phone Calls to Support Patients With Atrial Fibrillation: Randomized Controlled Trial

Ritu Trivedi[1], PhD; Liliana Laranjo[1], PhD; Simone Marschner[1], PhD; Aravinda Thiagalingam[1,2], PhD; Stuart Thomas[1,2], PhD; Saurabh Kumar[1,2], PhD; Tim Shaw[1,3*], PhD; Clara K Chow[1,2*], PhD

[1]Westmead Applied Research Centre, Faculty of Medicine and Health, The University of Sydney, Level 5, Block K, Westmead Hospital, Hawkesbury Road, Westmead, Australia

[2]Cardiology Department, Westmead Hospital, Westmead, Australia

[3]Charles Perkins Centre, Faculty of Medicine and Health, The University of Sydney, Camperdown, Australia

[*]these authors contributed equally

**Corresponding Author:**
Liliana Laranjo, PhD
Westmead Applied Research Centre, Faculty of Medicine and Health, The University of Sydney, Level 5, Block K, Westmead Hospital, Hawkesbury Road, Westmead, Australia

## Abstract

**Background:** Patient education and self-management support are critical for atrial fibrillation (AF) management. Conversational artificial intelligence (AI) has the potential to provide interactive and personalized support, but has not been evaluated in patients with AF.

**Objective:** This study aimed to evaluate the feasibility of a conversational AI intervention to support patients with AF postdischarge.

**Methods:** This was a single-blinded, 4:1-parallel–randomized controlled trial with process evaluation of feasibility and engagement. The primary outcome was the change in Atrial Fibrillation Effect on Quality-of-Life (AFEQT) questionnaire total score between groups. Patients with AF (18 y and older) were recruited postdischarge from Westmead Hospital cardiology services and randomized to receive either the intervention or usual care. The 6-month intervention consisted of fully automated conversational AI phone calls (with speech recognition and natural language processing) that regularly assessed patient health and symptoms and provided self-management support and education. These phone calls were supplemented with an online survey (sent via text message or email) containing replicated call content when participants could not be reached after 3 call attempts. If participant responses were concerning (eg, poor overall health, low medication confidence, and high symptom burden), they would be followed up with an ad hoc phone call and directed to clinical care if required. A semipersonalized education website was also available as part of the intervention, and participants were encouraged weekly (nudges delivered via text messages or emails) to visit it.

**Results:** A total of 103 patients (mean age, 63.7 y, SD 11.2 y; n=72, 70% male) were randomized (82 to the intervention); the target sample size was 385. The difference in the AFEQT total score was nonsignificant (adjusted mean difference 2.08, 95% CI −7.79 to 11.96; *P*=.46). An exploratory prepost comparison revealed an improvement in total AFEQT score in the intervention group only (baseline: 69.9, 95% CI 64.4 to 75.5; 6 months: 79.9, 95% CI 74.9 to 84.8; *P*=.01). Participants completed 4 of 7 outreaches on average, and 88.4% (304/344) of completed outreaches were reported as useful.

**Conclusions:** This proof-of-concept study demonstrates the feasibility of conversational AI in supporting patients with chronic conditions postdischarge. Intervention participants had improvement in their atrial fibrillation quality of life, though the forced shortening of the evaluation was unable to demonstrate a significant difference between groups.

**Trial Registration:** Australian New Zealand Clinical Trials Registry ACTRN12621000174886; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12621000174886

**International Registered Report Identifier (IRRID):** RR2-10.2196/34470

***(JMIR Cardio 2025;9:e64326)*** doi:10.2196/64326

**KEYWORDS**

atrial fibrillation; quality of life; natural language processing; self-management; digital health; randomized controlled trial; conversational artificial intelligence; feasibility; phone calls

## Introduction

The increase in atrial fibrillation (AF) prevalence is a global public health concern. AF presents the health system with various challenges: its rapidly increasing patient population [1], the multifaceted care required to manage patients and prevent outcomes of stroke and mortality [2], and the significant costs associated with hospitalizations [3]. Guidelines suggest a digitally enabled integrated approach to AF management, involving a multidisciplinary team to provide patient-centered care and support patient self-management (eg, lifestyle behavior change and medication adherence) [2].

Existing trials of digital interventions to support AF self-management have found mixed results, with some evidence of improvements in quality of life (QoL), knowledge, medication adherence, and clinical outcomes (composite outcome comprising stroke or thromboembolism, all-cause death, and rehospitalization) compared to usual care [4-6]. Most digital interventions for patients with AF have primarily been delivered through mobile apps, and many have reported low user engagement [7-9].

Conversational technologies now offer interactive ways of providing education and self-management support to patients [10]. These technologies can simulate human conversations through text or speech in an accessible and personalized manner, with users reporting high satisfaction. Recent voice-based conversational technologies leverage artificial intelligence (AI; including speech recognition and natural language processing) to facilitate more engaging and human-like dialogues [11,12]. There is limited literature available on the efficacy of conversation-AI interventions, and no studies have been conducted in an AF population [11-14].

The aim of this proof-of-concept randomized controlled trial (RCT) was to evaluate an intervention comprising conversational AI automated phone calls, text messages, and emails, and an educational website to better support patients with self-managing their AF [15]. The Coordinating Health Care With Artificial Intelligence–Supported Technology for Patients With Atrial Fibrillation (CHAT-AF) trial assessed the impact of this conversational AI intervention on Atrial Fibrillation Specific Quality of Life (AF-QoL), as well as evaluated its feasibility (engagement and perceived usefulness).

## Methods

### Study Design

CHAT-AF was a single-blinded parallel RCT, with a 4:1 allocation chosen to optimize process evaluation [15]. Trial registration with the Australian and New Zealand Clinical Trials Registry was submitted on November 25, 2020, but COVID-19 delays led to the trial only being registered on February 18, 2021 (85 d post initial submission; ACTRN12621000174886), and at this time, there were 22 participants enrolled.

### Patient Population

Participants were recruited from inpatient and outpatient cardiology services at Westmead Hospital. English-speaking adult patients with documented AF who had a mobile phone were eligible. Pregnant women and those participating in another clinical trial focused on providing AF education were excluded. Original plans to conduct a multicenter trial did not proceed after news of the acquisition of the technology partner by another organization (details below).

### Randomization and Masking

Randomization was 4:1 (intervention: control), stratified by sex, in blocks of 5. The allocation sequence was incorporated into REDCap (Research Electronic Data Capture; Vanderbilt University) [16] with access groups enabled, where data analysts were blinded. Participants, care providers, and research assistants were not blinded. Participants were randomized to receive usual care or the conversational AI intervention.

### Study Procedures

All participants completed study assessments via electronic surveys at baseline (in hospital or sent a link via text messages or email), 3-, and 6-months (via a link sent by text message or email). Baseline demographics and medical history were collected by participant self-report or by electronic medical records review (Multimedia Appendix 1). Questionnaires to assess primary and secondary outcomes were also completed electronically, either in person (baseline—during initial hospital visit) or sent via text message or email (3- and 6-month follow-up).

### Ethical Considerations

The human and research ethics committee at Western Sydney Local Health District (2020/ETH02546) approved this study. Informed consent, either written or over-the-phone consent, was obtained from all study participants. Data were collected and stored on secure servers accessible to approved study personnel only. Minimal participant data were provided to the technology partner via a secure file transfer protocol to enable delivery of the intervention phone calls, text messages, and emails. All data were deidentified for data analysis and publication. No compensation was offered to participants.

### Intervention (AF-Support)

The CHAT-AF intervention design and development have been previously described in detail elsewhere [15]. In summary, it consisted of 7 outreaches (“digital visits”) and a semipersonalized education website, which was available as part of the 6-month digital intervention (Multimedia Appendix 1). The technology in the intervention was provided by HMS (Health Management Systems, Inc).

The outreaches were delivered via fully automated conversational AI phone calls (with speech recognition and natural language processing capabilities) and supplemented with an online survey (a personalized link sent via text message or email) when participants could not be reached after 3 call attempts. Two main components underpinned the conversational AI in the automated calls [17]: (1) a speech recognition engine that recognized voice responses and translated them into text, and (2) natural language processing that identified the semantic and syntactic elements from user utterances, progressing the flow of the call depending on patient answers, in a decision tree

format. Given the proprietary nature of the software and the acquisition of the technology partner by another company, we were unable to obtain details about the network architecture, speech recognition, natural language processing capabilities, and other aspects related to the AI models in this intervention. During the phone calls, patients received AF education (eg, lifestyle information on diet and physical activity, importance of general practitioner [GP] follow-up, medication adherence, alcohol intake, blood pressure control, stroke, sleep apnea, and AF procedures) and were required to verbally respond to risk assessment queries (eg, overall health status, GP follow-up, AF symptoms and impact on daily life, medication confidence, and adherence). Certain patient responses to risk assessment questions would trigger alerts leading to an escalation pathway with clinician support, where needed. For example, if patients reported poor overall health status or significant impact of AF symptoms on daily life, these alerts would be actioned within 24 to 48 hours through a phone call by the researcher and additional clinical follow-up if required.

The 7 digital outreaches were also accompanied by weekly text messages or emails (depending on preferences) to a personalized link to an educational website that was tailored based on baseline characteristics (smoking status, alcohol consumption, hypertension diagnosis, and anticoagulant or warfarin prescription). This website contained AF-related information in the form of videos, texts, and images, as well as external links to online resources from reputable sources (eg, Heart Foundation and National Prescribing Service).

### Modified Delivery of Intervention and Premature Study Stopping

At the end of July 2021, we stopped recruitment as the delivery of the intervention was interrupted due to unforeseeable circumstances involving the acquisition of the technology partner by another organization. The trial steering committee made a decision to continue the trial to ensure the full 6-month program was delivered to all enrolled participants, by enabling feasible delivery of intervention content through text messages and emails. The technology partner had notified the research team in advance, allowing for the opportunity to develop an alternative approach and pre-emptively notify participants of the change. Further, comprehensive reports on each participant's intervention completion status were made available, which allowed for a more seamless transition period. At this point, 82 intervention participants were recruited, and of these, 20 had received all outreaches, with the remaining 62 being at differing stages of the intervention timeline. All participants had received at least 2 of 7 outreaches via the automated calls before premature study completion. The survey (delivered via REDCap) contained replicated content and questions as asked in the phone calls but required participants to click their responses. The hope had been to find another technology partner that could deliver the intervention according to our specifications, but we were unable to do this, and limited by the remaining budget, we stopped the study to report findings. These changes were planned, reviewed, and approved by the trial steering committee. This trial is reported according to the CONSERVE (CONSORT [Consolidated Standards of Reporting Trials] and SPIRIT [Standard Protocol Items: Recommendations for Interventional Trials] Extension for RCTs Revised in Extenuating Circumstances) statement [18].

### Outcomes

The primary outcome was change in AF-QoL assessed as Atrial Fibrillation Effect on Quality-of-Life (AFEQT) [19] total score at 6 months. AFEQT total score (0 - 100) is an average of subscales (symptom, daily activity, or treatment), with higher scores indicating better QoL.

Secondary outcomes included AFEQT subscales, AF knowledge [20], patient activation [21], patient assessment of care quality and self-management support [22], self-reported lifestyle behaviors, medication adherence (days of missed doses in prescription medications over the past week), health care service use (GP or cardiologist visits, emergency department presentation or hospitalization, and ablation or cardioversion procedure), and health outcomes (stroke or myocardial infarction). All outcomes were assessed at baseline and 6 months; AF-QoL (AFEQT) was also assessed at 3 months. A detailed list of outcomes, methods, and time of data collection is provided online (Multimedia Appendix 1).

Process evaluation outcomes for the intervention group included: outreach completion and perceived usefulness, and individual engagement (Multimedia Appendix 1). Outreach completion was calculated as the number of outreaches with at least half of the questions answered divided by the number of participants that received the outreach. Outreach perceived usefulness was defined as the number of individuals who answered, "yes" to the question "Did you find the information in this call/survey helpful?" and was divided by the number of participants who attempted the outreach (answered at least 1 question). Individual engagement was calculated as the number of outreaches completed by the participant and categorized, where 4 or more completed outreaches (of 7) were considered as "higher engagement." Metrics for engagement with the educational website were also explored.

### Statistical Analysis

A sample size of 385 was required to detect a between-group difference of 7 in the total score of the AFEQT questionnaire with 80% power ($\alpha$=.05; SD=19), accounting for a dropout rate of 10% [19,23]. The current study was limited in detecting differences in primary outcome as we were only able to recruit 27% (103 participants) of the intended sample size.

Analyses were prespecified in a statistical analysis plan (Multimedia Appendix 2 [15,19-22,24]) and were conducted according to intention-to-treat principles. Analyses were performed using R statistical software (version 4.1.2; R Project for Statistical Computing). Outcomes were analyzed using either ANCOVA or logistic regression, adjusting for baseline variables. All tests were 2-tailed, a $P$ value of <.05 was considered significant, and odds ratios were reported with 95% CIs. Normally distributed continuous variables were expressed as mean and SD. Nonnormally distributed variables were expressed as the median and IQR.

A univariate logistic regression analysis to predict higher engagement was conducted with covariates of age, gender,

ethnicity, education, type of AF, time since AF diagnosis, and $CHA_2DS_2$-VASC score (congestive heart failure, hypertension, age 75 y and older [2 points], diabetes, stroke [2 points], vascular disease, aged 65 to 74 years, and sex category [female]; calculated as a sum, where 1-point or 2-points [where indicated] is given when aforementioned characteristics are present) [24].

## Results

### Sample Characteristics

Between January and July 2021, we enrolled 103 participants (82 intervention and 21 control; Figure 1). We lost 16 intervention participants and 3 controls to follow-up (18.4%), and the primary outcome analysis included 66 intervention participants and 19 controls. The follow-up period was from July 2021 to April 2022. Mean age was 63.7 (SD 11.2) years and 69.9% were males (Table 1). The majority were nonuniversity graduates (75.5%), of non-Caucasian ethnicity (24.5%), and had paroxysmal AF (76.7%). Detailed characteristics are provided online (Multimedia Appendix 1).

**Figure 1.** Flowchart of participants included in the CHAT-AF randomized controlled trial. CHAT-AF: Coordinating Health Care With Artificial Intelligence–Supported Technology for Patients With Atrial Fibrillation.

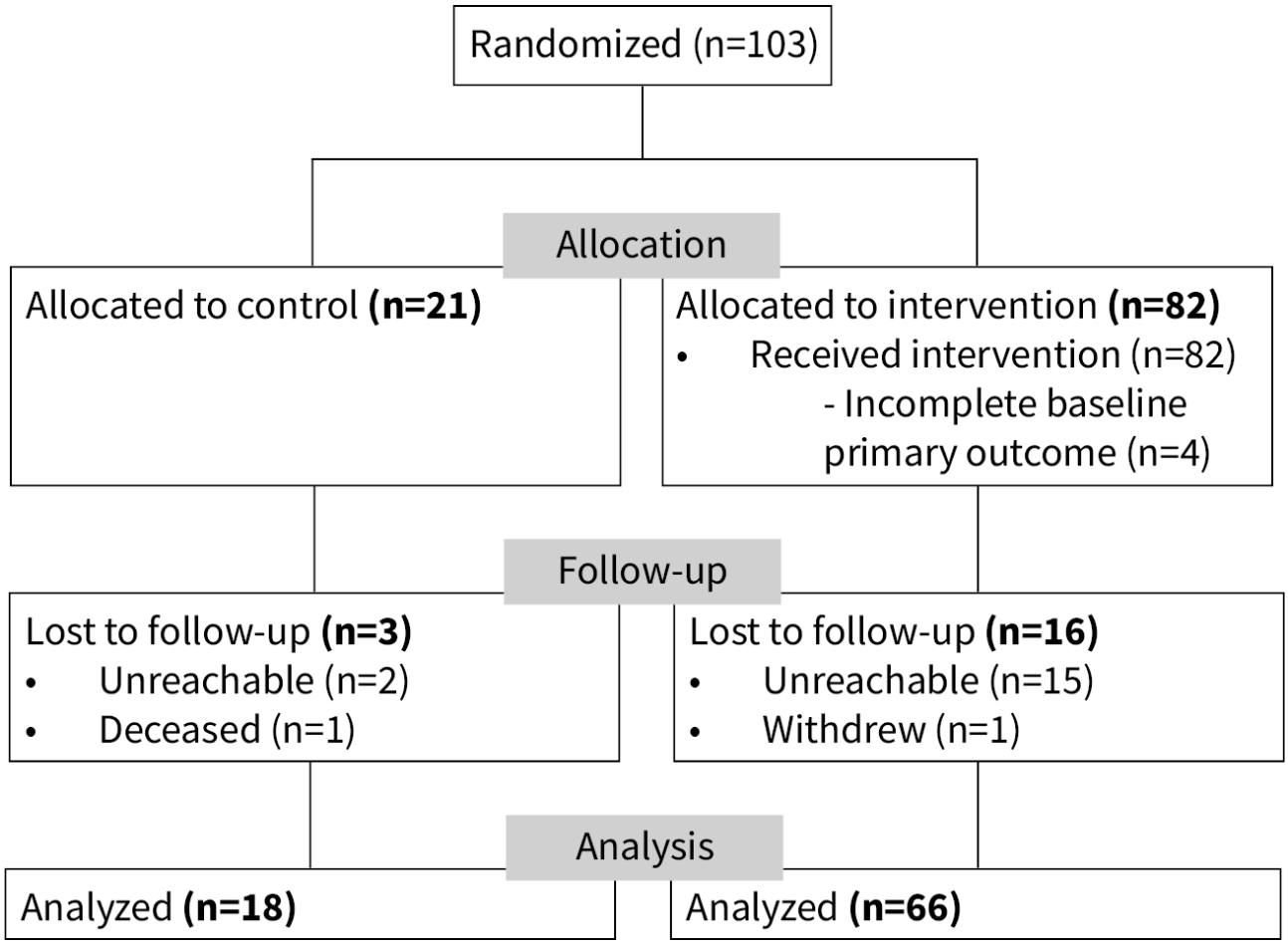


**Table .** Participant baseline characteristics (N=103).

| | | Control (n=21) | Intervention (n=82) | Total (n=103) |
|---|---|---|---|---|
| Age[a] (years), mean (SD) | | 63.0 (12.1) | 63.8 (11.0) | 63.7 (11.2) |
| BMI[a] (kg/m$^3$), mean (SD) | | 30.9 (5.1) | 31.6 (6.8) | 31.4 (6.5) |
| Sex[a], n (%) | | | | |
| | Male | 14 (66.7) | 58 (70.7) | 72 (69.9) |
| | Female | 7 (33.3) | 24 (29.3) | 31 (30.1) |
| Blood pressure[a] (mm Hg), mean (SD) | | | | |
| | Systolic | 130.7 (17.2) | 132.1 (18.0) | 131.8 (17.8) |
| | Diastolic | 78.8 (14.6) | 77.6 (13.4) | 77.8 (13.6) |
| Ethnicity, n (%) | | | | |
| | Non-Caucasian | 4 (19) | 21 (25.9) | 25 (24.5) |
| Education, n (%) | | | | |
| | Nonuniversity graduate | 13 (61.9) | 64 (79) | 77 (75.5) |
| Annual household income in AUS $[b], n (%) | | | | |
| | 0-31,199 | 4 (28.6) | 9 (16.7) | 13 (19.1) |
| | 31,200-77,999 | 3 (21.4) | 20 (37) | 23 (33.8) |
| | 78,000+ | 7 (50) | 25 (46.3) | 32 (47.1) |
| Smoking status, n (%) | | | | |
| | Never smoked | 8 (38.1) | 39 (50) | 47 (47.5) |
| | Current smoker | 0 (0) | 7 (9) | 7 (7.1) |
| | Ex-smoker | 13 (61.9) | 32 (41) | 45 (45.5) |
| Type of AF[ac] (most recent), n (%) | | | | |
| | Paroxysmal | 15 (71.4) | 64 (78) | 79 (76.7) |
| | Persistent | 5 (23.8) | 17 (20.7) | 22 (21.4) |
| | Permanent | 1 (4.8) | 0 (0) | 1 (1) |
| | Unspecified | 0 (0) | 1 (1.2) | 1 (1) |
| Time since initial AF diagnosis (years), n (%) | | | | |
| | <5 | 13 (61.9) | 44 (56.4) | 57 (57.6) |
| | ≥5 | 8 (38.1) | 34 (43.6) | 42 (42.4) |
| Medical history[a], n (%) | | | | |
| | Hypertension | 12 (57.1) | 54 (65.9) | 66 (64.1) |
| | Hyperlipidemia | 10 (47.6) | 39 (47.6) | 49 (47.6) |
| | Heart failure | 7 (33.3) | 23 (28) | 30 (29.1) |
| | Vascular disease | 8 (38.1) | 24 (29.3) | 32 (31.1) |
| | Stroke | 2 (9.5) | 6 (7.3) | 8 (7.8) |
| | Diabetes type 2 | 5 (23.8) | 15 (18.3) | 20 (19.4) |
| | Obstructive sleep apnea | 6 (28.6) | 33 (40.2) | 39 (37.9) |
| $CHA_2DS_2$-VASC[d] score, mean (SD) | | 2.7 (2.2) | 2.5 (1.6) | 2.6 (1.7) |
| Medications[a], n (%) | | | | |
| | Antiarrhythmic | 14 (66.7) | 63 (76.8) | 77 (74.8) |
| | Anticoagulation | 15 (71.4) | 68 (82.9) | 83 (80.6) |

| | | Control (n=21) | Intervention (n=82) | Total (n=103) |
|---|---|---|---|---|
| | Statin | 12 (57.1) | 46 (56.1) | 58 (56.3) |
| | Angiotensin-converting enzyme inhibitor | 3 (14.3) | 16 (19.5) | 19 (18.4) |
| | Angiotensin receptor blocker | 4 (19) | 28 (34.1) | 32 (31.1) |
| | Calcium channel blocker | 5 (23.8) | 13 (15.9) | 18 (17.5) |
| | Neprilysin inhibitor | 1 (4.8) | 7 (8.5) | 8 (7.8) |

[a]Information collected by clinical investigators from the electronic medical record.

[b]The average conversion rate during the study was AUS $1=US $0.73.

[c]AF: atrial fibrillation.

[d]$CHA_2DS_2$-VASC: congestive heart failure, hypertension, age 75 years and older (doubled), diabetes, stroke (doubled), vascular disease, aged 65 to 74 years, and sex category (female).

## Primary and Secondary Outcomes

No significant difference was observed between groups in the primary outcome of AFEQT total score (2.08, 95% CI −7.79 to 11.96; *P*=.46; Figure 2 and Table 2). There were 18.4% missing primary outcome data at 6 months; however, we found no evidence for differences in missingness based on age (grouped by 65 y or older), gender, or ethnicity (grouped by Caucasian or non-Caucasian). Prespecified sensitivity analyses were conducted; both imputation of 3-month AFEQT total score carried forward and baseline imputation revealed similar results to the primary analysis.

**Figure 2.** Box plot of AFEQT total and subscale scores. AFEQT scores for intervention (n=66) and control (n=18) groups at B and 6M. Total AFEQT score is an average of symptoms, activity, and treatment subscales (does not include satisfaction). * Significant difference of prepost Wilcoxon 2-tailed *t* test. AFEQT: Atrial Fibrillation Effect on Quality-of-Life; B: baseline; 6M: 6 months.

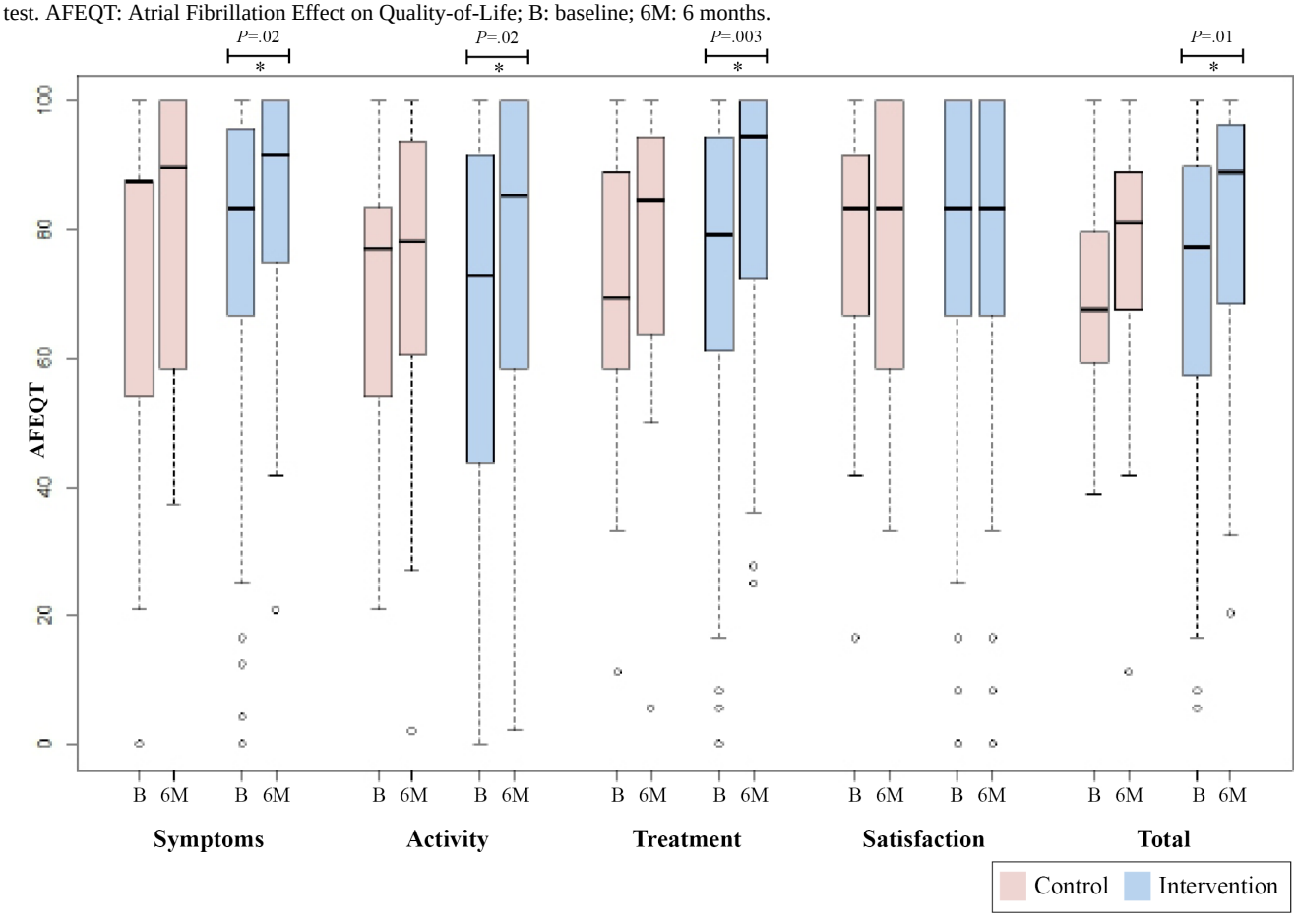


**Table .** Primary outcome—AFEQT[a].

| AFEQT | Control | | | Intervention | | | | |
|---|---|---|---|---|---|---|---|---|
| | Baseline, mean (SD; n=21) | Follow-up, mean (SD; n=18) | Change, mean (95% CI) | Baseline, mean (SD; n=78) | Follow-up, mean (SD; n=66) | Change, mean (95% CI) | Difference, mean (95% CI) | *P* value[b] |
| Total score | 70.3 (17.0) | 74.9 (22.7) | 5.0 (−6.1 to 6.0) | 69.9 (25.0) | 79.9 (20.5) | 7.1 (2.9 to 11.3) | 2.08 (−7.79 to 11.96) | .46 |
| Symptom | 72.0 (29.1) | 79.9 (20.9) | 6.9 (−9.8 to 23.6) | 74.3 (26.8) | 83.1 (19.2) | 5.8 (0.5 to 11.1) | −1.14 (−14.43 to 12.15) | .69 |
| Activity | 69.4 (21.9) | 70.6 (27.2) | 1.3 (−9.1 to 11.6) | 65.6 (30.6) | 75.0 (26.6) | 7.2 (1.3 to 13.1) | 5.92 (−6.59 to 18.44) | .35 |
| Treatment | 70.2 (24.4) | 77.5 (24.5) | 8.6 (−5.7 to 23.0) | 72.7 (26.0) | 84.3 (20.3) | 7.7 (3.5 to 12.0) | −0.90 (−11.85 to 10.05) | .54 |
| Satisfaction | 76.6 (21.8) | 76.9 (24.0) | 1.4 (−13.9 to 16.7) | 76.7 (25.6) | 80.1 (24.1) | 2.3 (−4.3 to 8.9) | 0.88 (−14.05 to 15.82) | .69 |

[a]AFEQT: Atrial Fibrillation Effect on Quality-of-Life.

[b]Adjusted analysis consisted of an analysis of covariance test, adjusting for baseline level to estimate the difference between groups at 6 months. Atrial Fibrillation Effect on Quality-of-Life questionnaire scores range from 0 - 100 (higher scores associated with better health-related quality of life). The total Atrial Fibrillation Effect on Quality-of-Life score is an average of all subscales (total score, symptom, activity, and treatment), excluding satisfaction.

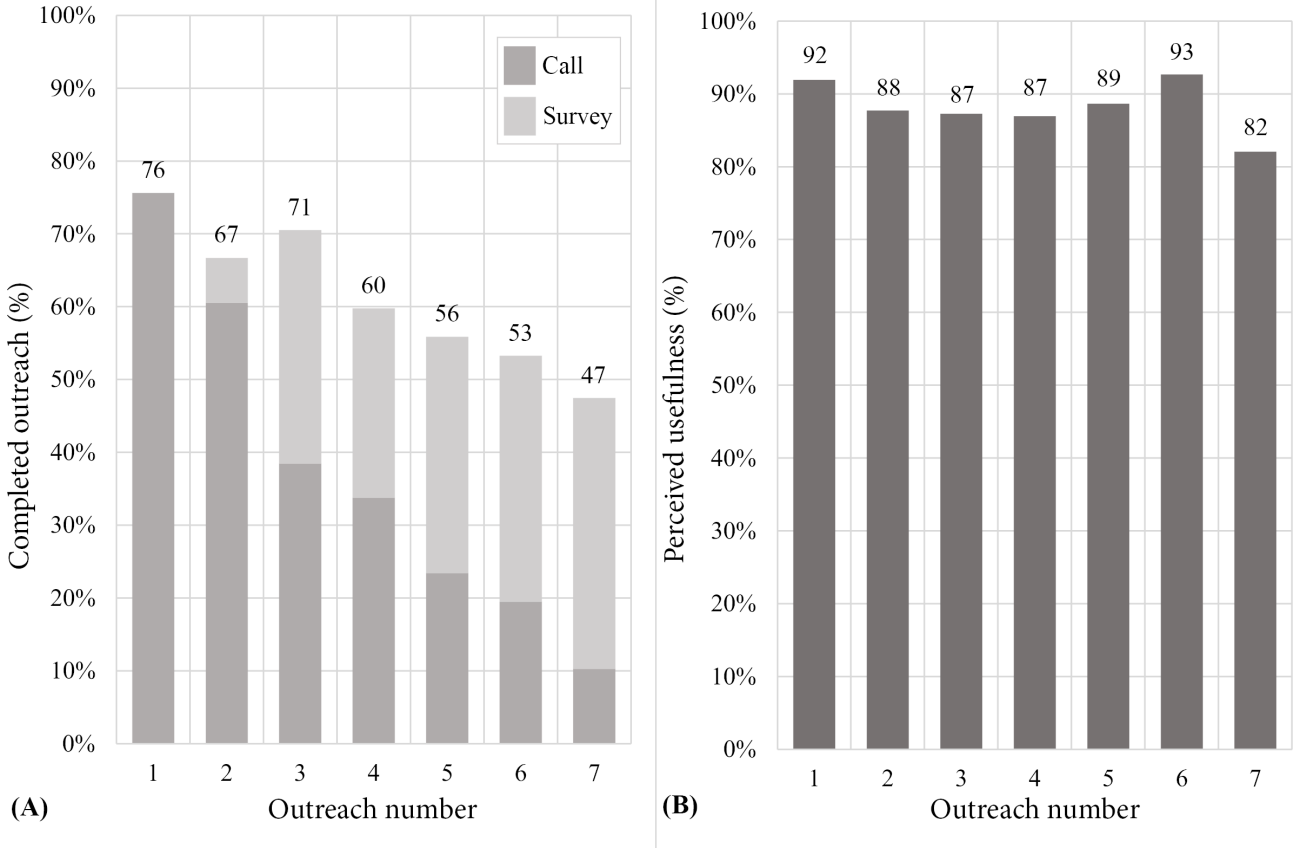
No difference was observed between groups in the change in AFEQT subscale scores (Figure 2, Table 2). Additional exploratory analyses revealed an improvement in AFEQT total score postintervention in the intervention group (baseline: 69.9, 95% CI 64.4 to 75.5; 6-months: 79.9, 95% CI 74.9 to 84.8; *P*=.01), with no improvements in the control group (Figure 2). Within-group differences revealed the intervention group improved in most AFEQT subscales postintervention (symptom, daily activity, treatment, *P* values<.05), with no improvements seen in the control group (Figure 2, Table 2). No significant differences were observed in AFEQT total and subscale scores from baseline to 3 months between groups (Multimedia Appendix 1).

No significant differences were observed in secondary outcomes of knowledge (AF knowledge), patient activation (Patient Activation Measure), patient assessment of care quality and self-management (Patient Assessment of Chronic Illness Care), or lifestyle behaviors (exercise, fruit and vegetable intake, alcohol intake, and smoking) at 6 months (Multimedia Appendix 1). No significant difference was observed between groups in the proportion that were adherent to medications, visited a GP or cardiologist, visited the emergency department, or were hospitalized and had an AF procedure (Multimedia Appendix 1). A total of 3% of intervention participants had a stroke or myocardial infarct in the previous 6 months compared to no controls (Multimedia Appendix 1).

## Process Evaluation

A total of 338 of 550 outreaches were delivered via the conversational AI calls, and of these, 226 (66.9%) were completed (original delivery; Multimedia Appendix 1). The remaining 212 outreaches were delivered only by a survey tool, and of these, 112 (52.8%) were completed (modified delivery). The completion rate of the first outreach was 75.6% (62 calls and 0 surveys), and this dropped to 47.4% (8 calls and 29 surveys) by the final outreach, with an average completion rate across the 7 outreaches of 61.5% (Figure 3). On average, each participant completed 4.12 of 7 outreaches. A total of 51 participants (62.2%) had higher engagement (completed 4 or more outreaches), and there were no demographic variables influencing this outcome (age, gender, ethnicity, education, type of AF, time since AF diagnosis, and $CHA_2DS_2$-VASC). Most participants (56.1%) visited the educational website at least once, and the mean number of visits was 5.54 times. The most visited topic was general information about AF (138 visits), which included videos narrated by a local cardiologist. In terms of perceived usefulness, 88.4% of completed outreach was reported as useful (Figure 3) and this was similar in the original (89.1%) and the modified delivery (87.6%). Detailed process evaluation results are provided online (Multimedia Appendix 1).

**Figure 3.** CHAT-AF intervention (A) outreaches completion and (B) perceived usefulness. Outreaches 1 to 7 occurred at 24 - 48 hours, 2 weeks, 1 month, 2 months, 3 months, 4 months, and 5 months posthospital discharge, respectively. CHAT-AF: Coordinating Health Care With Artificial Intelligence–Supported Technology for Patients With Atrial Fibrillation.




## Discussion

### Principal Results

The CHAT-AF study provides a proof-of-concept with initial data on the efficacy of a novel digital health follow-up strategy for patients with AF that leverages conversational AI. On average, participants completed 4 of 7 outreaches, and 88.4% of completed outreaches were reported as useful, suggesting this approach to delivering care is feasible and could enable chronic disease services to follow up with patients at scale with fewer in-person and staffed visits. Due to challenges with the technology partner acquisition, this study had only recruited 27% of its intended sample size and was subsequently underpowered for its primary outcome. The trial did not demonstrate a significant difference in the primary outcome between intervention and control; however, there was a benefit suggested by an exploratory prepost comparison that showed overall AF-QoL improved at 6 months in the intervention group from baseline, with no similar improvement observed in the control group.

### Comparison With Prior Work

Digital support programs show promise in improving QoL for patients with AF, but more robust studies are needed to determine effectiveness. Despite this study's limited power to determine efficacy, other trials have reported improvements in AF-QoL, knowledge, medication adherence, and clinical outcomes, with similar digital health interventions (without conversational AI) that delivered a combination of health education, monitoring, and self-management support for patients with AF [4-6,25]. However, the paucity of literature in this space has been confined to mobile apps [4,6,25-27], text messaging [26,28], and web-based platforms [5,8,9], many of which have reported limited engagement. The current intervention was unique in its approach to interactively engage with patients through conversational technology. A comparable study by Guhl et al 2020 [4] involved an embodied conversational agent that consisted of a virtual avatar displaying verbal and nonverbal gestures to deliver education and heart rhythm monitoring support to patients with AF [4]. Participants interacted with the conversational agent 18 times over 30 days and had improved medication adherence and AF-QoL compared with usual care [4]. Notably, this intervention did not include an AI component and required users to respond to queries by clicking on prespecified options in the mobile app, rather than using speech recognition technology, which was used in CHAT-AF to facilitate more natural dialogue and better engagement.

AI-enabled conversational technologies allow for more human-like interactions and have shown promise in other populations. Existing trials of conversational AI have demonstrated improvements in insulin adherence and glycemic control in patients with type 2 diabetes [12], medication adherence in patients with hypertension and diabetes [11], and symptoms of depression and anxiety in college students [13,14]. These studies reported good engagement over a 2-week period, where college students exchanged 283 messages with a chatbot [14], and undertook 12 check-ins with another chatbot [13]

regarding their mental health. Another study that used a voice-based conversational AI interface, resembling the phone calls in the current intervention, found patients with type 2 diabetes logged daily insulin use and fasting blood glucose levels almost every day for about 4 months [12]. Similar to the phone call–based delivery mode used in the current study, another study delivered cognitive behavioral therapy for pain management using voice-based conversational-AI phone calls and reported that 87.8% of calls were completed [29]. In our study, the transition from phone call to solely survey-based delivery may have resulted in completion rates (67%) lower than reported in the aforementioned study. Overall, evidence suggests conversational AI technologies can successfully engage and support patients with chronic disease management, but additional studies are needed to evaluate long-term engagement.

The deeper level of human-technology interaction allowed by conversational AI seems a key factor in achieving higher user engagement [12-14,30-32]. The unplanned change in intervention delivery in the current study offered an important opportunity to compare engagement between different communication modes. Interestingly, the shift from conversational AI phone calls to surveys resulted in a drop in overall outreach completion rates (from 66.9% to 52.8%). A hypothesized explanation may be the stronger appeal of the human-like interactions provided by these technologies and their ability to facilitate natural dialogues with users. This is in line with existing work indicating the possibility of a relationship between humans and nonhuman agents in the context of health [12-14]. As digital health interventions increasingly incorporate elements that make them interactive, adaptive, persuasive, and personalized, they become more able to reproduce elements of a therapeutic relationship and can better engage and support patients in their health journey [33].

The early stopping of the trial due to suspended delivery of the intervention by the technology partner posed challenges that are important to consider. For the trial steering committee, it was key that enrolled participants received the entirety of the 6-month program, to satisfy the duty of care for enrolled patients as well as to ensure fidelity in delivery of all educational content in the program. The decision to use REDCap [16] surveys when the automated phone calls had to be stopped offered a satisfactory solution to these concerns and also provided the opportunity to observe feasibility measures such as engagement and perceived usefulness rates between these 2 modes of delivery while the other elements were unchanged [30,31]. An attenuating factor for early stopping was that the trial steering committee was forewarned of when the company was going to withdraw services, and that allowed for adequate time to develop an alternative solution and streamline the transition for participants. Other articles have reported early trial stopping due to technology partner withdrawal of services [34]. Open communication and goal alignment have been proposed as key to achieving solutions that can benefit all parties involved and bridge the gap in “academia-industry” relationships [35,36]. Stronger partnership between sectors is increasingly needed as collaboration between technological, research, and clinical expertise is paramount to successfully implement digital health solutions.

## Implications

The growing prevalence of cardiovascular diseases such as AF, which require ongoing management to prevent frequent hospitalization, puts a significant toll on health care resources. Interventions such as CHAT-AF have the potential to provide support at scale, with a risk management system that allows patients at risk of deteriorating to be identified, prioritized, and managed appropriately. Using phone calls to provide patient support has advantages over other mobile technologies (eg, apps and wearable devices) as it is not dependent on internet connectivity, phone models, or operating systems. Digital interventions such as CHAT-AF could be used to provide support to patients with AF in the community, including remote and rural areas where patients have traditionally had poorer access to health care services [37]. As demonstrated in this proof-of-concept study, the engagement, perceived usefulness, and initial suggestive findings of AF-QoL improvement argue that this technology should be further examined in larger RCTs.

## Limitations

This study has several limitations that need to be considered. This study was limited in not achieving its target sample size and multicenter reach due to premature study completion, and therefore, careful consideration needs to be made when interpreting the results as this study was underpowered. This also meant some participants on the intervention arm were unable to complete their conversational AI “visits” later into the planned follow-up period, which is likely to have contributed to the ability to evaluate the effectiveness of the intervention on outcomes. The authors are in the process of conducting a large multicenter RCT evaluation of an optimized version of the current intervention, to address the power, sample size, and interpretation limitations of this study (registered with the Australian and New Zealand Clinical Trials Registry, ACTRN12623000850673). Moreover, except for baseline characteristics (medical history), all outcome data were self-reported. There was a difference in attrition between the intervention (16/82) and control (3/21) groups, which may have impacted the primary analysis, and while there was no evidence of bias for missingness, caution needs to be taken in interpretation due to the limitations in lack of power and changes in intervention delivery. This study’s population was younger (mean age of 63.7 y) than the average patients with AF population in Australia (mean age of 75 y) [38], which may have occurred due to the higher likelihood of younger patients’ comfort and interest in participating in technology-based research [35]. Perceived usefulness was assessed as the final question of each outreach and therefore results are based on individuals that completed the particular outreach and its final question—this may not be reflective of overall intervention usefulness, as it does not reflect outreaches not completed by participants (344 of 550 outreaches were completed by participants and were rated in terms of their perceived usefulness, of these, 304 of the 344, 88.4%, were assessed positively based on usefulness). However, to capture the perspectives of a diverse group of the intervention cohort, we conducted qualitative interviews using purposive sampling techniques to ensure we captured the experiences of those with varying levels of engagement [38,39]. Furthermore, the study

was delivered in English and included predominately Caucasian participants, highlighting the need for further work to validate this technology in larger trials with more diverse and representative patient populations.

### Conclusions

This study found that a conversational AI follow-up program for patients with AF improved AF-QoL postintervention (but not compared with usual care). As the burden of AF continues to grow, novel technologies that can interact with patients and support them in their care journey will be needed, and digital health can provide this at a scalable level. However, larger-scale RCTs and implementation studies are needed to determine the effectiveness of conversational AI in improving AF outcomes.

## Acknowledgments

The authors would also like to acknowledge Ms Emma Charlston, who was involved in project management, Dr Harry Klimis for his involvement in this study and intervention design, and Ms Haeri Min, who independently validated the primary outcome analysis. Additionally, we would like to thank the technology partner (HMS [Health Management Systems, Inc]) and team members for their efforts in the delivery of the intervention. This work was supported by Digital Health CRC Limited (DHCRC). DHCRC is funded under the Commonwealth's Cooperative Research Centres Program.

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

RT, LL, SM, AT, ST, SK, TS, CKC, Ms Emma Charlston, and Dr Harry Klimis conceptualized this study. RT did the data curation, formal analysis, investigation, worked on the software and visualization, and writing of the original draft. CKC and TS acquired the funding. RT, LL, SM, AT, ST, SK, TS, CKC, and Ms Emma Charlston handled the methodology of this study. RT, LL, and Ms Emma Charlston were responsible for project administration. LL, TS, and CKC supervised this study. Ms Haeri Min validated the data. RT, LL, SM, AT, ST, SK, TS, and CKC reviewed and edited the writing.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Intervention details, outcome measures, and detailed analyses.
[PDF File, 3126 KB - cardio_v9i1e64326_app1.pdf ]

Multimedia Appendix 2
Statistical analysis plan.
[PDF File, 706 KB - cardio_v9i1e64326_app2.pdf ]

Checklist 1
CONSORT-EHEALTH checklist. CONSORT: Consolidated Standards of Reporting Trials.
[PDF File, 2320 KB - cardio_v9i1e64326_app3.pdf ]

## References

1. Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation 2014 Feb 25;129(8):837-847. [doi: 10.1161/CIRCULATIONAHA.113.005119] [Medline: 24345399]
2. Hindricks G, Potpara T, Dagres N, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. Eur Heart J 2021 Feb 1;42(5):373-498. [doi: 10.1093/eurheartj/ehaa612] [Medline: 32860505]
3. Burdett P, Lip GYH. Atrial fibrillation in the UK: predicting costs of an emerging epidemic recognizing and forecasting the cost drivers of atrial fibrillation-related costs. Eur Heart J Qual Care Clin Outcomes 2022 Mar 2;8(2):187-194. [doi: 10.1093/ehjqcco/qcaa093] [Medline: 33346822]
4. Guhl E, Althouse AD, Pusateri AM, et al. The atrial fibrillation health literacy information technology trial: pilot trial of a mobile health app for atrial fibrillation. JMIR Cardio 2020 Sep 4;4(1):e17162. [doi: 10.2196/17162] [Medline: 32886070]

5. Toscos T, Coupe A, Wagner S, et al. Engaging patients in atrial fibrillation management via digital health technology: the impact of tailored messaging. J Innov Card Rhythm Manag 2020 Aug;11(8):4209-4217. [doi: 10.19102/icrm.2020.110802] [Medline: 32874747]
6. Guo Y, Lane DA, Wang L, et al. Mobile health technology to improve care for patients with atrial fibrillation. J Am Coll Cardiol 2020 Apr 7;75(13):1523-1534. [doi: 10.1016/j.jacc.2020.01.052] [Medline: 32241367]
7. Goldenthal IL, Sciacca RR, Riga T, et al. Recurrent atrial fibrillation/flutter detection after ablation or cardioversion using the AliveCor KardiaMobile device: iHEART results. J Cardiovasc Electrophysiol 2019 Nov;30(11):2220-2228. [doi: 10.1111/jce.14160] [Medline: 31507001]
8. Desteghe L, Germeys J, Vijgen J, et al. Effectiveness and usability of an online tailored education platform for atrial fibrillation patients undergoing a direct current cardioversion or pulmonary vein isolation. Int J Cardiol 2018 Dec 1;272:123-129. [doi: 10.1016/j.ijcard.2018.07.065] [Medline: 30049498]
9. Chen YC, Roebuck AE, Sami A, Ersin Ö, Mirro MJ. The use of electronic personal health records to improve medication adherence and patient engagement: a randomized study of non-valvular atrial fibrillation patients. J Innov Card Rhythm Manag 2017 Aug;8(8):2804-2813. [doi: 10.19102/icrm.2017.080803] [Medline: 32494465]
10. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 1;25(9):1248-1258. [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
11. Kassavou A, Mirzaei V, Brimicombe J, et al. A highly tailored text and voice messaging intervention to improve medication adherence in patients with either or both hypertension and type 2 diabetes in a UK primary care setting: feasibility randomized controlled trial of clinical effectiveness. J Med Internet Res 2020 May 19;22(5):e16629. [doi: 10.2196/16629] [Medline: 32427113]
12. Nayak A, Vakili S, Nayak K, et al. Use of voice-based conversational artificial intelligence for basal insulin prescription management among patients with type 2 diabetes: a randomized clinical trial. JAMA Netw Open 2023 Dec 1;6(12):e2340232. [doi: 10.1001/jamanetworkopen.2023.40232] [Medline: 38039007]
13. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 6;4(2):e19. [doi: 10.2196/mental.7785] [Medline: 28588005]
14. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health 2018 Dec 13;5(4):e64. [doi: 10.2196/mental.9782] [Medline: 30545815]
15. Laranjo L, Shaw T, Trivedi R, et al. Coordinating health care with artificial intelligence-supported technology for patients with atrial fibrillation: protocol for a randomized controlled trial. JMIR Res Protoc 2022 Apr 13;11(4):e34470. [doi: 10.2196/34470] [Medline: 35416784]
16. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
17. Ram A, Prasad R, Khatri C, Venkatesh A, Gabriel R, Liu Q, et al. Conversational AI: the science behind the Alexa prize. arXiv. Preprint posted online on Jan 11, 2018. [doi: 10.48550/arXiv.1801.03604]
18. Orkin AM, Gill PJ, Ghersi D, et al. Guidelines for reporting trial protocols and completed trials modified due to the COVID-19 pandemic and other extenuating circumstances. JAMA 2021 Jul 20;326(3):257. [doi: 10.1001/jama.2021.9941]
19. Spertus J, Dorian P, Bubien R, et al. Development and validation of the Atrial Fibrillation Effect on Quality-of-Life (AFEQT) Questionnaire in patients with atrial fibrillation. Circ Arrhythm Electrophysiol 2011 Feb;4(1):15-25. [doi: 10.1161/CIRCEP.110.958033] [Medline: 21160035]
20. Hendriks JML, Crijns H, Tieleman RG, Vrijhoef HJM. The atrial fibrillation knowledge scale: development, validation and results. Int J Cardiol 2013 Sep 30;168(2):1422-1428. [doi: 10.1016/j.ijcard.2012.12.047] [Medline: 23305860]
21. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the Patient Activation Measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930. [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
22. Glasgow RE, Wagner EH, Schaefer J, Mahoney LD, Reid RJ, Greene SM. Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). Med Care 2005 May;43(5):436-444. [doi: 10.1097/01.mlr.0000160375.47920.8c] [Medline: 15838407]
23. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11. [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
24. Mark DB, Anstrom KJ, Sheng S, et al. Effect of catheter ablation vs medical therapy on quality of life among patients with atrial fibrillation. JAMA 2019 Apr 2;321(13):1275. [doi: 10.1001/jama.2019.0692]
25. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach. Chest 2010 Feb;137(2):263-272. [doi: 10.1378/chest.09-1584]
26. Guo Y, Chen Y, Lane DA, Liu L, Wang Y, Lip GYH. Mobile health technology for atrial fibrillation management integrating decision support, education, and patient involvement: mAF app trial. Am J Med 2017 Dec;130(12):1388-1396. [doi: 10.1016/j.amjmed.2017.07.003] [Medline: 28847546]

27. Turakhia M, Sundaram V, Smith SN, et al. Efficacy of a centralized, blended electronic, and human intervention to improve direct oral anticoagulant adherence: Smartphones to Improve Rivaroxaban Adherence in Atrial Fibrillation (SmartADHERE) a randomized clinical trial. Am Heart J 2021 Jul;237:68-78. [doi: 10.1016/j.ahj.2021.02.023] [Medline: 33676886]
28. Desteghe L, Kluts K, Vijgen J, et al. The health buddies app as a novel tool to improve adherence and knowledge in atrial fibrillation patients: a pilot study. JMIR mHealth uHealth 2017 Jul 19;5(7):e98. [doi: 10.2196/mhealth.7420] [Medline: 28724512]
29. Caceres BA, Hickey KT, Bakken SB, et al. Mobile ECG monitoring and health-related quality of life in patients with atrial fibrillation: findings from the iHEART study. J Cardiovasc Nurs 2020;35(4):327. [doi: 10.1097/JCN.0000000000000646] [Medline: 32015256]
30. Piette JD, Newman S, Krein SL, et al. Patient-centered pain care using artificial intelligence and mobile health tools: a randomized comparative effectiveness trial. JAMA Intern Med 2022 Sep 1;182(9):975-983. [doi: 10.1001/jamainternmed.2022.3178] [Medline: 35939288]
31. Short C, Rebar A, Plotnikoff R, Vandelanotte C. Designing engaging online behaviour change interventions: a proposed model of user engagement. Eur Psychol 2015;17(1):32-38 [FREE Full text]
32. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions. JMIR Form Res 2019 Oct 10;3(4):e14052. [doi: 10.2196/14052] [Medline: 31603427]
33. Yardley L, Spring BJ, Riper H, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
34. Olivier CB, Middleton SK, Purington N, et al. Why digital health trials can fail: lessons learned from a randomized trial of health coaching and virtual cardiac rehabilitation. Cardiovasc Digit Health J 2021 Apr;2(2):101-108. [doi: 10.1016/j.cvdhj.2021.01.003] [Medline: 35265897]
35. Austin D, May J, Andrade J, Jones R. Delivering digital health: the barriers and facilitators to university-industry collaboration. Health Policy Technol 2021 Mar;10(1):104-110. [doi: 10.1016/j.hlpt.2020.10.003]
36. Liu C, Shao S, Liu C, Bennett GG, Prvu Bettger J, Yan LL. Academia-industry digital health collaborations: a cross-cultural analysis of barriers and facilitators. Digit Health 2019;5:2055207619878627. [doi: 10.1177/2055207619878627] [Medline: 31632684]
37. Essien UR, Kornej J, Johnson AE, Schulson LB, Benjamin EJ, Magnani JW. Social determinants of atrial fibrillation. Nat Rev Cardiol 2021 Nov;18(11):763-773. [doi: 10.1038/s41569-021-00561-0] [Medline: 34079095]
38. Wong CX, Brooks AG, Cheng YH, et al. Atrial fibrillation in Indigenous and non-Indigenous Australians: a cross-sectional study. BMJ Open 2014 Oct 24;4(10):e006242. [doi: 10.1136/bmjopen-2014-006242] [Medline: 25344486]
39. Trivedi R, Shaw T, Sheahen B, Chow CK, Laranjo L. Patient perspectives on conversational artificial intelligence for atrial fibrillation self-management: qualitative analysis. J Med Internet Res 2025 Mar 12;27:e64325. [doi: 10.2196/64325] [Medline: 40073398]


## Abbreviations

**AF:** atrial fibrillation
**AF-QoL:** Atrial Fibrillation Specific Quality of Life
**AFEQT:** Atrial Fibrillation Effect on Quality-of-Life
**AI:** artificial intelligence
**$CHA_2DS_2$-VASC:** congestive heart failure, hypertension, age 75 years and older (doubled), diabetes, stroke (doubled), vascular disease, aged 65 to 74 years, and sex category (female)
**CHAT-AF:** Coordinating Health Care With Artificial Intelligence–Supported Technology for Patients With Atrial Fibrillation
**CONSERVE:** CONSORT (Consolidated Standards of Reporting Trials) and SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) Extension for RCTs (Randomized Controlled Trials) Revised in Extenuating Circumstances
**CONSORT :** Consolidated Standards of Reporting Trials
**GP:** general practitioner
**HMS:** Health Management Systems, Inc
**QoL:** Quality of Life
**RCT:** randomized controlled trial
**REDCap:** Research Electronic Data Capture
**SPIRIT :** Standard Protocol Items: Recommendations for Interventional Trials

*Edited by A Coristine; submitted 17.07.24; peer-reviewed by A Olivella, M Wright, ST Arasteh; revised version received 04.04.25; accepted 04.04.25; published 19.08.25.*

*Please cite as:*
*Trivedi R, Laranjo L, Marschner S, Thiagalingam A, Thomas S, Kumar S, Shaw T, Chow CK*
*Conversational AI Phone Calls to Support Patients With Atrial Fibrillation: Randomized Controlled Trial*
*JMIR Cardio 2025;9:e64326*
*URL: https://cardio.jmir.org/2025/1/e64326*
*doi:10.2196/64326*

# Predicting Atrial Fibrillation Relapse Using Bayesian Networks: Explainable AI Approach

João Miguel Alves[1,2], MSc; Daniel Matos[3], MD; Tiago Martins[1,2], MSc; Diogo Cavaco[3], MD; Pedro Carmo[3], MD; Pedro Galvão[3], MD; Francisco Moscoso Costa[3], MD; Francisco Morgado[3], MD; António Miguel Ferreira[3], MD; Pedro Freitas[3], MD; Cláudia Camila Dias[1,2], PhD; Pedro Pereira Rodrigues[1,2*], PhD; Pedro Adragão[3*], PhD

[1]Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Rua Dr Plácido da Costa, Porto, Portugal

[2]CINTESIS @ RISE – Center for Health Technology and Services Research, Porto, Portugal

[3]Cardiology and Electrophysiology Department, Hospital de Santa Cruz, Centro Hospitalar de Lisboa Ocidental, Carnaxide, Portugal

[*]these authors contributed equally

**Corresponding Author:**
João Miguel Alves, MSc
Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Rua Dr Plácido da Costa, Porto, Portugal

## Abstract

**Background:** Atrial fibrillation (AF) is a prevalent arrhythmia associated with significant morbidity and mortality. Despite advancements in ablation techniques, predicting recurrence of AF remains a challenge, necessitating reliable models to identify patients at risk of relapse. Traditional scoring systems often lack applicability in diverse clinical settings and may not incorporate the latest evidence-based factors influencing AF outcomes. This study aims to develop an explainable artificial intelligence model using Bayesian networks to predict AF relapse postablation, leveraging on easily obtainable clinical variables.

**Objective:** This study aims to investigate the effectiveness of Bayesian networks as a predictive tool for AF relapse following a percutaneous pulmonary vein isolation (PVI) procedure. The objectives include evaluating the model's performance using various clinical predictors, assessing its adaptability to incorporate new risk factors, and determining its potential to enhance clinical decision-making in the management of AF.

**Methods:** This study analyzed data from 480 patients with symptomatic drug-refractory AF who underwent percutaneous PVI. To predict AF relapse following the procedure, an explainable artificial intelligence model based on Bayesian networks was developed. The model used a variable number of clinical predictors, including age, sex, smoking status, preablation AF type, left atrial volume, epicardial fat, obstructive sleep apnea, and BMI. The predictive performance of the model was evaluated using the area under the receiver operating characteristic curve (AUC-ROC) metrics across different configurations of predictors (5, 6, and 7 variables). Validation was conducted through four distinct sampling techniques to ensure robustness and reliability of the predictions.

**Results:** The Bayesian network model demonstrated promising predictive performance for AF relapse. Using 5 predictors (age, sex, smoking, preablation AF type, and obstructive sleep apnea), the model achieved an AUC-ROC of 0.661 (95% CI 0.603 - 0.718). Incorporating additional predictors improved performance, with a 6-predictor model (adding BMI) achieving an AUC-ROC of 0.703 (95% CI 0.652 - 0.753) and a 7-predictor model (adding left atrial volume and epicardial fat) achieving an AUC-ROC of 0.752 (95% CI 0.701 - 0.800). These results indicate that the model can effectively estimate the risk of AF relapse using readily available clinical variables. Notably, the model maintained acceptable diagnostic accuracy even in scenarios where some predictive features were missing, highlighting its adaptability and potential use in real-world clinical settings.

**Conclusions:** The developed Bayesian network model provides a reliable and interpretable tool for predicting AF relapse in patients undergoing percutaneous PVI. By using easily accessible clinical variables, presenting acceptable diagnostic accuracy, and showing adaptability to incorporate new medical knowledge over time, the model demonstrates a flexibility and robustness that makes it suitable for real-world clinical scenarios.

***(JMIR Cardio 2025;9:e59380)*** doi:10.2196/59380

**KEYWORDS**

artificial intelligence; atrial fibrillation; Bayesian networks; clinical decision-making; machine learning; prognostic models

## Introduction

Atrial fibrillation (AF), the most common sustained cardiac arrhythmia [1], poses significant challenges in the clinical management and prediction of disease progression. Currently, the ATLAS score [2] provides a reliable risk estimate to predict the rate of AF recurrence after a pulmonary vein isolation (PVI) procedure. However, it suffers from typical limitations of clinical scores, such as the use of a fixed number of independent variables for the prediction of a single dependent variable, its static nature, and its inability to be adjusted as new knowledge becomes available. All these issues can be addressed by artificial intelligence (AI) models based on machine learning algorithms, which can learn from available data, be quickly updated with new data, and perform complex calculations in a short time.

In recent years, such machine learning techniques have emerged as powerful tools in various medical domains, including cardiology [3,4]. There have been some recent successful attempts to develop AI models to predict the recurrence of AF after ablation procedure. However, despite the good performance of those models, they either lack the explainability required to allow their acceptance by health care professionals [5,6], or share the same limitations of medical scores discussed above [7]. In fact, although many physicians have recognized that AI models may be useful both for diagnosis and prognosis in medical practice, many authors raise legitimate questions about the lack of explainability of some AI models [8,9].

Bayesian networks, despite being still poorly adopted in health care, have gained popularity as clinical decision support models in medicine due to their ability to handle complex problems with causal dependencies, integrate both data and domain knowledge, provide an interpretable graphical structure, and support both diagnostic and prognostic reasoning [10]. In addition, these models can be updated with new medical knowledge, enabling the incorporation of novel risk factors and advancements in the field of arrhythmology. This adaptability and scalability make Bayesian networks a promising tool for decision-making in medicine and long-term monitoring of patients with AF.

This study aims to address key research gaps in the prediction of AF relapse by developing a more reliable and adaptable predictive model based on Bayesian networks. Traditional medical scoring systems are limited by their reliance on a fixed set of independent variables, which reduces their generalizability across diverse patient populations. In addition, many existing AI models for AF prediction lack the necessary explainability required to foster trust and acceptance among health care professionals. To bridge these gaps, this study makes several significant contributions. First, it introduces a novel explainable AI model based on Bayesian networks, which allows for the calculation of conditional probabilities tailored to individual patient profiles, thus enhancing both the interpretability of the predictions and their clinical acceptance. Second, the study overcomes the limitations of traditional scoring systems by offering a dynamic and adaptable model that can incorporate new risk factors and learn from evolving patient data, thereby improving predictive accuracy over time. Third, the proposed model demonstrates flexibility and robustness, making it suitable for real-world clinical scenarios where incomplete data may be present. Finally, by integrating this model into clinical decision support systems, the study has the potential to enhance decision-making processes and improve patient outcomes in the management of AF. In this work, we investigate the use of Bayesian networks to predict AF relapse before a percutaneous PVI procedure and evaluate its potential as a valuable clinical tool, with the primary aim of improving clinical decision-making and patient care.

## Methods

### Study Population

All consecutive patients with symptomatic drug-refractory AF undergoing cardiac computed tomography (CT) before percutaneous PVI at Hospital Santa Cruz (Carnaxide, Portugal) between November 2015 and July 2019 were included in an observational registry used for this retrospective study. Patients with moderate or severe valvular heart disease, left atrial thrombus, abnormal thyroid function, or contraindication to anticoagulation were excluded. Baseline demographic and clinical characteristics, including age, sex, height, weight, and presence of hypertension, diabetes, smoking, and known coronary artery disease, were recorded for all patients. AF was categorized as paroxysmal if it self-terminated in less than 7 days, persistent if episodes lasted ≥7 days or required cardioversion, or long-standing persistent if AF was maintained for more than 12 months.

### PVI Protocol

PVI was guided by electroanatomical mapping, using either NavX (St Jude Medical) or CARTO (Biosense Webster) systems. The right femoral vein was used as the preferred vascular access, through which three catheter electrodes were introduced: (1) a decapolar catheter, advanced through the coronary sinus; (2) a variable circular mapping catheter, placed in the pulmonary veins (PVs); and (3) an irrigated contact force-sensing ablation catheter. Left atrial access was established by a transseptal puncture. Radiofrequency ablation was performed more than 5 mm from the PV ostia, with continuous lesions enclosing the left and right pairs of PVs. The treatment was considered successful if complete electrophysiological PVI was achieved. When required, electrical cardioversion was performed at the end of the procedure. Oral anticoagulation was resumed 6 hours after the ablation, maintained for 6 months, and then withdrawn or continued according to CHA2DS2-VASc criteria. Generally, class I/III antiarrhythmic drugs were maintained in all patients for the first 3 months after the procedure and then withdrawn if there was no AF recurrence. A proton pump inhibitor was also prescribed for the first month after the ablation.

### Study End Point and Patient Follow-Up

The study end point was AF recurrence, defined as symptomatic or documented AF or other atrial arrhythmias, after a 3-month blanking period. Symptomatic AF was defined as the presence of symptoms considered to be likely due to AF episodes. Documented AF was defined by the presence of at least one

episode of AF lasting more than 30 seconds in an ECG, 24-hour Holter monitoring, or event-loop recording. The follow-up protocol comprised outpatient visits with 12-lead ECG and 24-hour Holter monitoring at the assistant physicians' discretion (typically at 6 and 12 months, and yearly thereafter). Patients were encouraged to contact the department if they experienced symptoms of AF recurrence. Whenever clinical records were insufficient, a structured telephonic interview was conducted. Patients who were kept on antiarrhythmic drugs after the third month of follow-up were not considered as failed ablation.

## Population Characteristics

The analyzed sample comprised demographic and clinical data from 480 patients who underwent follow-up after the PVI procedure described above. The cohort included 295 (61.5%) men and 185 (38.5%) women, with a mean age of 61.1 (SD 11.5) years. The median duration of the follow-up time of the patients was 392 (IQR 150 - 674) days. For the purpose of this study, all numeric variables in the dataset (including age, BMI, left atrial volume, and epicardial fat) were discretized into classes. Data characterization is shown in Table 1.

**Table .** Demographic and clinical characteristics of the patients included in the study.

| Characteristics | Total (N=480), n (%) | AF[a] relapse (n=166), n (%) | AF-free (n=314), n (%) |
|---|---|---|---|
| **Sex** | | | |
| Female | 185 (38.5) | 55 (33.1) | 130 (41.4) |
| Male | 295 (61.5) | 111 (66.9) | 184 (58.6) |
| **Age (years)** | | | |
| ≤45 | 57 (11.9) | 9 (5.4) | 48 (15.3) |
| 46 - 65 | 234 (48.8) | 84 (50.6) | 150 (47.8) |
| +65 | 189 (39.4) | 73 (44) | 116 (36.9) |
| Alcoholism | 25 (5.2) | 15 (9) | 10 (3.2) |
| Smoking | 135 (28.1) | 57 (34.3) | 78 (24.8) |
| Diabetes | 46 (9.6) | 16 (9.6) | 30 (9.6) |
| High blood pressure | 292 (60.8) | 105 (63.3) | 187 (59.6) |
| Obstructive sleep apnea | 50 (10.4) | 35 (21.1) | 15 (4.8) |
| **BMI** | | | |
| Normal weight | 151 (31.5) | 35 (21.1) | 116 (36.9) |
| Overweight | 218 (45.4) | 74 (44.6) | 144 (45.9) |
| Obese | 111 (23.1) | 57 (34.3) | 54 (17.2) |
| **Atrial fibrillation** | | | |
| Paroxysmal | 374 (77.9) | 98 (59) | 276 (87.9) |
| Persistent | 106 (22.1) | 68 (41) | 38 (12.1) |
| **Left atrium volume[b] ($ml/m^2$)** | | | |
| [0 to 100] | 168 (35) | 39 (23.5) | 129 (41.1) |
| (100 to 125] | 172 (35.8) | 56 (33.7) | 116 (36.9) |
| (125 to inf) | 140 (29.2) | 71 (42.8) | 69 (22) |
| **Epicardial fat[b] ($cm^3$)** | | | |
| [0 to 2.7] | 162 (33.8) | 18 (10.8) | 144 (45.9) |
| (2.7 to 4.6] | 166 (34.6) | 48 (28.9) | 118 (37.6) |
| (4.6 to inf) | 152 (31.7) | 100 (60.2) | 52 (16.6) |

[a]AF: atrial fibrillation.

[b]Square brackets indicate that the end point is included in the range, and parentheses indicate that the end point is not included in the range.

The variable preablation AF type represents the type of AF identified in each patient before the ablation procedure, being coded either as paroxysmal or persistent. The variable sex is categorized as binary (female or male). All other binary variables such as alcoholism, smoking, diabetes, high blood pressure, and obstructive sleep apnea, were coded as logical (true or false), indicating the presence or absence of that condition.

The variable AF relapse represents the identification of postprocedural AF relapse in patients during follow-up

examinations, also coded as logical (true or false). It was targeted as the outcome variable for this study.

### Bayesian Network Model Training

#### *Network Structure*

Considering that Bayesian networks are probabilistic graphical models made to represent knowledge, we started by building our network structure primarily based on medical knowledge in this field. In a first step, we opted to include (whitelist) some of the most noteworthy known clinical relationships between features, such as (1) known risk factors for diseases expressed in the dataset, namely diabetes, high blood pressure (HBP), and obstructive sleep apnea (OSA); and (2) known predictive features of AF relapse, such as the ATLAS score features (age, sex, smoking, persistent AF and left atrial volume), as well as epicardial fat [11,12] and OSA [13,14], as suggested by recent medical literature.

In the second step, we explored additional potential relationships between features that could improve model fit and better explain the observed data through data-driven inference. To achieve this, we applied a score-based structure learning method, using the Bayesian Information Criterion (BIC) [15] as the scoring metric to be optimized. The optimization of the BIC score was performed using a hill-climbing algorithm [16]. This approach allowed us to learn the remaining structure of the network, resulting in a model that aligns with current medical knowledge while effectively capturing the relationships between the variables.

#### *Model Fitting*

After the network structure was defined, a model could be set to learn the conditional probabilities among all related features. The parameters of the Bayesian network were thus fit given the previously learned structure and the available data, by means of a Bayesian posterior estimator with a uniform before. With the model fitted in this fashion, it was now possible to use the model to compute the estimated probability that a given patient has AF relapse given her clinical characteristics, for example, the model can be asked “based on the available data, what is the probability that a patient has AF relapse knowing that she is female,+65 years old and non-smoking.” Further examples of computed conditional probabilities for AF relapse based on patients’ conditions are presented in the *Results* section.

### Model Validation

Model validation was executed by out-of-sample testing to assess the predictive performance of the model on unseen data, as follows: from the full dataset, a random sample was taken to be used as training data for the model. This sample was used to train a conditional probabilities model, as previously described. Following that, the remaining observations that were not included in the training set were used as a test set, upon which the model predictions were tested. For this testing step, we used the model to compute the conditional probability of AF relapse for each patient in the test set, and stored the prediction results for each tested observation. This process was cyclically repeated multiple times until each observation had been used for testing at least 30 times. Finally, the calculated probability of AF relapse for each patient was assumed to be the average of all estimated probabilities for that patient. We then compared the average predicted probability with the true observation of AF relapse for each patient, and measured the performance through the area under the receiver operating characteristic curve (AUC-ROC).

Regarding the sampling process at the beginning of each cycle, it is worth mentioning that the random samples for training the model were obtained through one of four different sampling processes: (1) bootstrapping, which on average uses 63.2% of the observations for training, or (2) hold-out, using fixed splitting ratios for the train and test of 80:20, (3) 90:10, and (4) 95:5, that is, with 80%, 90%, and 95% of the observations, respectively, being used for training the model, and the remaining proportion used for testing. With these processes, we aimed to assess the model’s ability to generalize for unknown data and achieve a good estimator for the generalization error.

This analysis was carried out using R (version 4.2.2; R Foundation for Statistical Computing) [17], with packages *bnlearn* [18] and *pROC* [19].

### Ethical Considerations

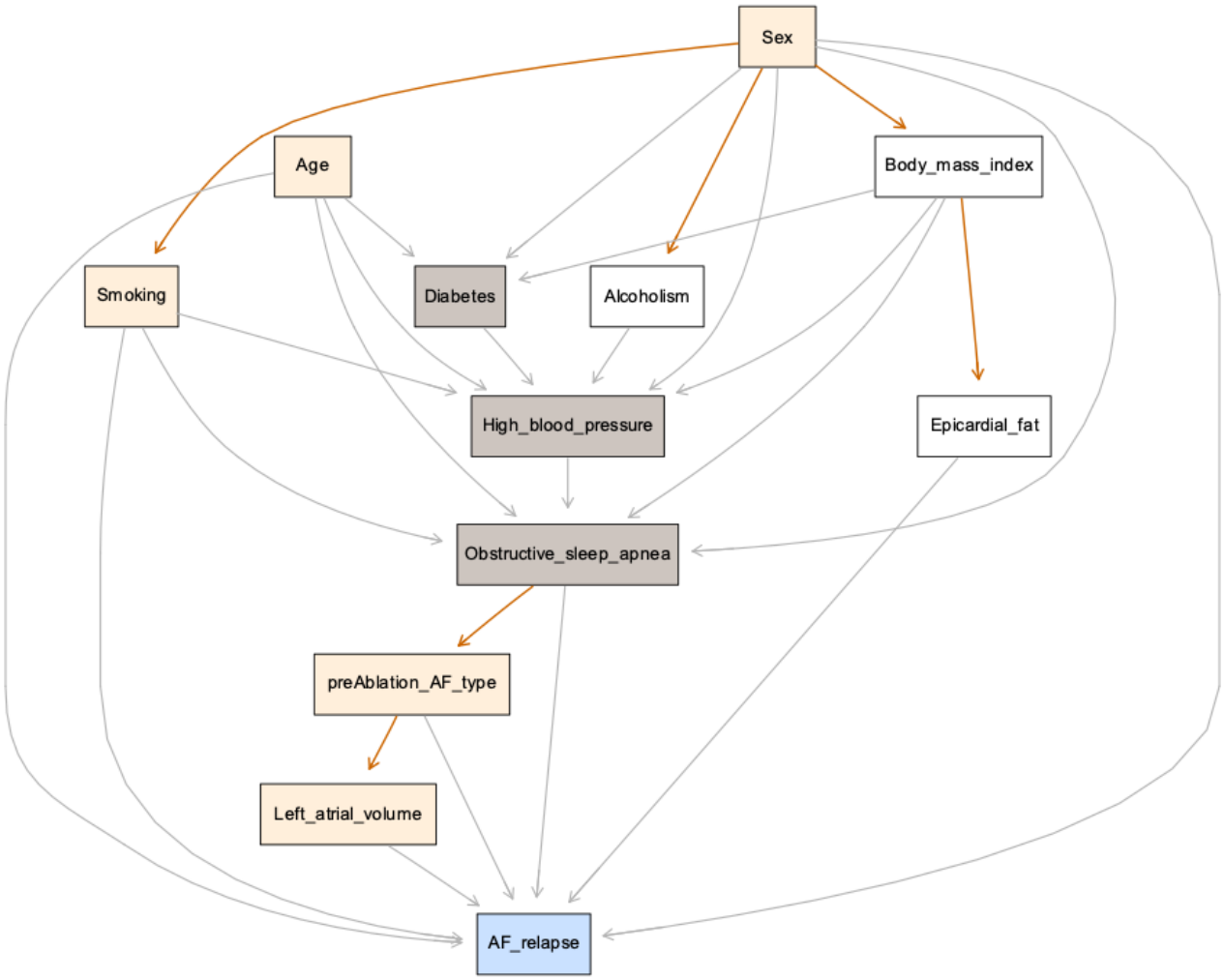
This study adheres to the ethical guidelines of the Declaration of Helsinki, including its later amendments. It has been approved by the Health Ethics Commission of the Western Lisbon Hospital Center, with the approval number 2117. All patients provided written informed consent before this study for both the procedure and the publication of any relevant data. Patient confidentiality was maintained by removing any personally identifiable information from all data used in this study and its supplementary materials.

## Results

### Bayesian Network Structure

The Bayesian network structure defined by expert knowledge and inference from data is represented in Figure 1.

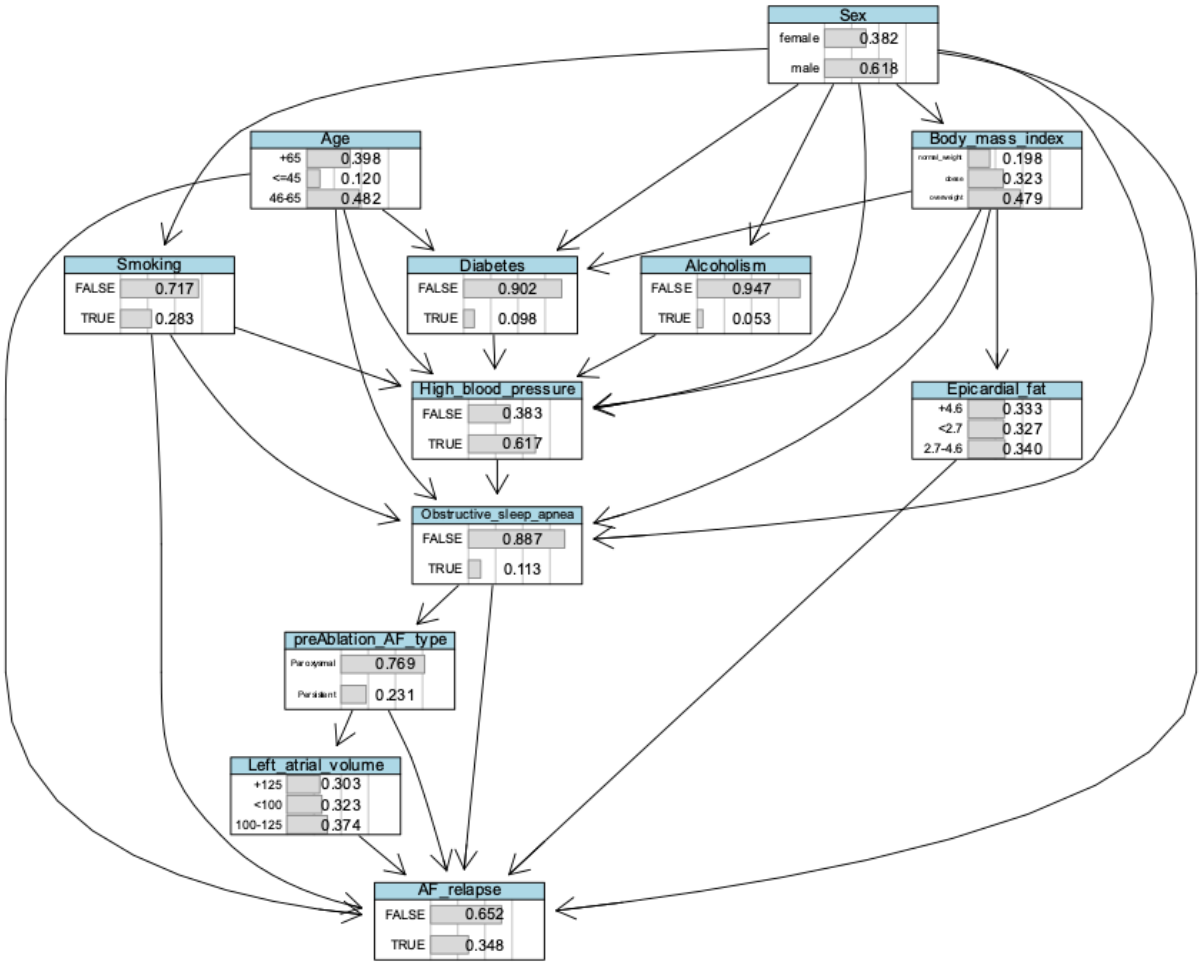
**Figure 1.** Bayesian network structure with nodes (boxes) representing the analyzed demographic and clinical variables. Grey nodes represent diseases with known associated risk factors, namely diabetes, high blood pressure, and obstructive sleep apnea. Beige nodes represent the 5 atrial fibrillation (AF) relapse predictors used by the ATLAS score, namely age, sex, smoking status, preablation AF type, and left atrial volume. The blue node highlights AF relapse as the outcome variable. The arcs (arrows) represent the direction of influence of variables. Grey arcs represent manually input relationships deriving from medical knowledge, ie, known risk factors. Orange colored arcs represent relationships discovered by the artificial intelligence algorithm, suggesting other meaningful relationships between variables.




As noted in this representation, the model suggests relationships that were not initially declared, such as BMI→Epicardial fat, OSA→preablation AF type, and preablation AF type→Left atrial volume. Furthermore, sex appears to be related to active smoking, alcoholism, and BMI. All these relationships are not surprising and are even supported by the current medical literature, thus providing a reasonable representation of clinical knowledge in this field. Regarding the outcome variable AF relapse, the model did not find any other relevant relations apart from those previously whitelisted.

An alternative representation of this network is exhibited in Figure 2, showing relative frequencies per class at each node.

**Figure 2.** Bayesian network structure with node-specific tables displaying relative frequencies per class at each node. AF: atrial fibrillation.




## Conditional Probability Calculation

With each trained model, we calculated the conditional probability of AF relapse for each patient in the test set, considering their reported clinical conditions. These probabilities were compared with the true values of AF relapse for each patient and plotted in a receiver operating characteristic (ROC) curve, with cutoff values for classification determined as those that maximize the Youden J statistic. We tested in turns 7, 5, or 6 predictive features, as explained in the sections to follow. For illustration purposes, Table 2 presents a few examples of different combinations of patients' conditions and their calculated conditional probability of AF relapse. These calculations were conducted for hypothetical patients, while considering as predictors all 7 parent nodes of AF relapse as represented in the network structure.

**Table .** Conditional probabilities of atrial fibrillation (AF) relapse for a sample of different combinations of hypothetical patients' conditions. Conditions are sorted from the most unlikely to experience AF relapse to the most likely to experience that outcome.

| Sex | Age (years) | Left atrium volume[a] (ml/m$^2$) | Smoking active | Persistent AF | Epicardial fat[a] (cm$^3$) | OSA[b] | Conditional probability of AF relapse, % (95% CI) |
|---|---|---|---|---|---|---|---|
| Male | ≤45 | [0 to 100] | False | Paroxysmal | [0 to 2.7] | False | 7.5 (1.8-13.2) |
| Male | 46 - 65 | (100 to 125] | False | Paroxysmal | [0 to 2.7] | False | 10.1 (6.3-13.8) |
| Female | ≤45 | [0 to 100] | False | Paroxysmal | [0 to 2.7] | False | 16.8 (7.4-26.1) |
| Male | 46 - 65 | (125 to inf) | False | Paroxysmal | (2.7 to 4.6] | False | 20.1 (14.3-26) |
| Male | +65 | (100 to 125] | True | Paroxysmal | (2.7 to 4.6] | False | 25.2 (17.3-33.1) |
| Male | 46 - 65 | (100 to 125] | True | Persistent | [0 to 2.7] | False | 33.2 (18.4-47.9) |
| Male | 46 - 65 | (100 to 125] | False | Paroxysmal | (4.6 to inf) | True | 33.3 (16.4-50.3) |
| Male | +65 | (125 to inf) | False | Paroxysmal | (2.7 to 4.6] | False | 33.3 (25.2-41.5) |
| Female | 46 - 65 | [0 to 100] | False | Paroxysmal | (2.7 to 4.6] | False | 40.1 (34-46.2) |
| Male | 46 - 65 | [0 to 100] | True | Paroxysmal | (4.6 to inf) | False | 50 (41.4-58.6) |
| Female | ≤45 | (100 to 125] | False | Paroxysmal | (4.6 to inf) | False | 50.1 (35.7-64.5) |
| Male | 46 - 65 | (100 to 125] | True | Paroxysmal | (4.6 to inf) | False | 66.3 (57.4-75.1) |
| Female | +65 | (125 to inf) | False | Persistent | (4.6 to inf) | False | 66.4 (53.8-78.9) |
| Male | +65 | (100 to 125] | False | Persistent | (4.6 to inf) | False | 66.4 (52.6-80.2) |
| Female | 46 - 65 | (100 to 125] | False | Paroxysmal | (4.6 to inf) | False | 66.5 (59.9-73.1) |
| Male | +65 | (125 to inf) | False | Paroxysmal | (4.6 to inf) | False | 71.5 (63.8-79.2) |
| Male | +65 | (125 to inf) | False | Persistent | (4.6 to inf) | False | 74.8 (63.3-86.4) |
| Male | 46 - 65 | (125 to inf) | True | Persistent | (4.6 to inf) | True | 74.9 (58.4-91.4) |

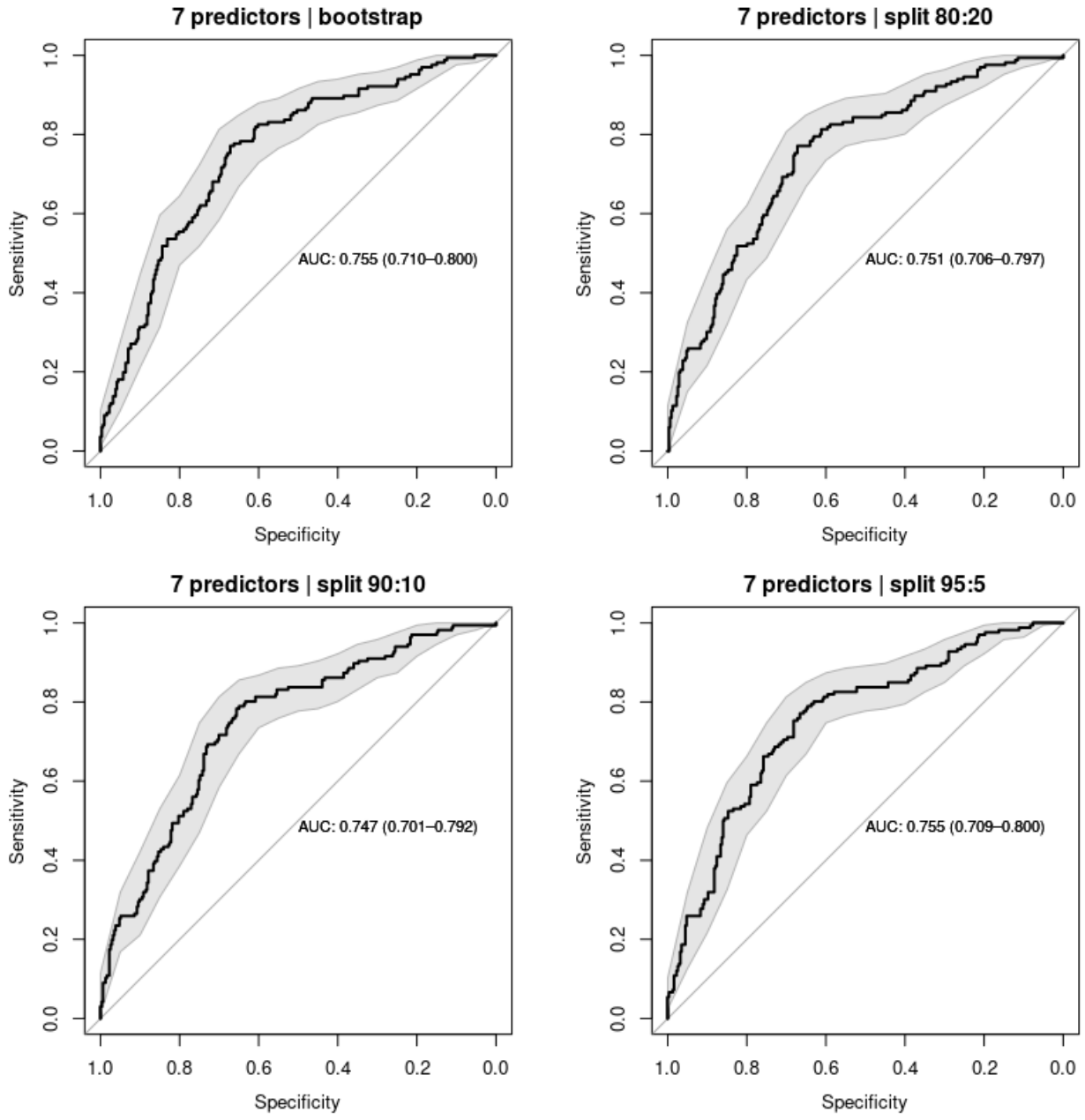
[a]Square brackets indicate that the end point is included in the range, and parentheses indicate that the end point is not included in the range.

[b]OSA: obstructive sleep apnea.

## The 7 Predictors

In the first stage, the calculation considered the clinical state of the patients for the 7 parent nodes of AF relapse represented in the network structure: age, sex, smoking, preablation AF type, left atrial volume, epicardial fat, and OSA. The performance of the model in classifying AF relapse with all parent nodes (7 predictors) was calculated to an average area under the curve (AUC) value of 0.752 (95% CI 0.701 - 0.800) for all sampling methods. ROC curves for each validation test are shown in Figure 3.

**Figure 3.** Receiver operating characteristic curves for all validation sampling methods applied to the model with 7 predictors: age, sex, smoking, preablation AF type, left atrial volume, epicardial fat, and obstructive sleep apnea. AUC values averaged 0.752 (95% CI 0.701 - 0.800). AUC: area under the curve.



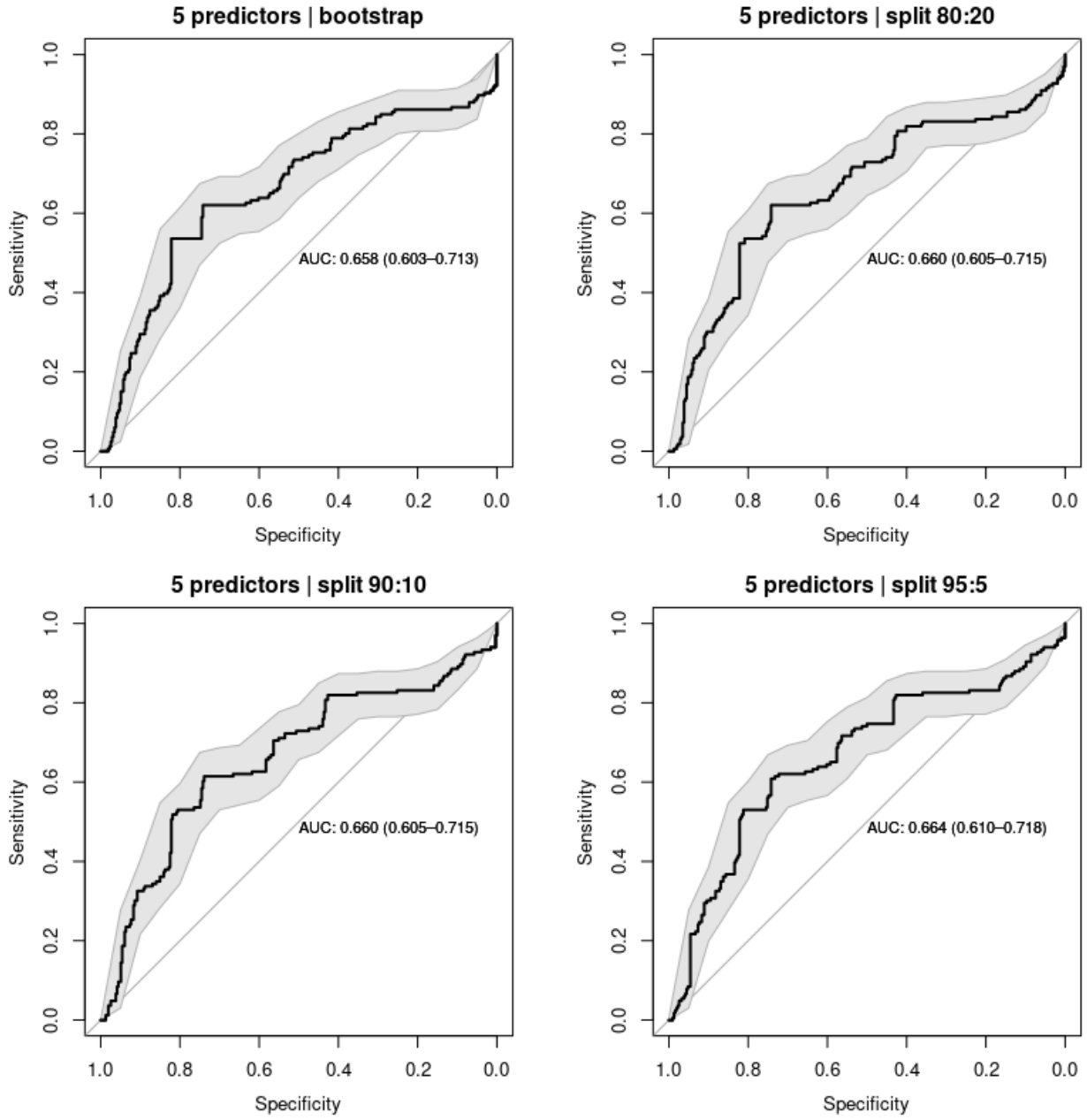

## The 5 Predictors

Out of the 7 predictive features used in the previous test, 2 are usually difficult to obtain: left atrial volume and epicardial fat. These 2 features are typically calculated by diagnostic imaging, which is not always performed for all patients. In some cases, the physician does not have access to those measurements, which frustrates the calculation of medical scores that require any of those values, as is the case with the ATLAS score.

The purpose of this test was to evaluate the performance of the model without these 2 features, thus simulating a frequent real-life scenario. As such, we calculated the conditional probability of AF relapse for each patient in the test set, considering only 5 of its parent nodes: age, sex, smoking, preablation AF type, and OSA. The remaining 2 parent nodes (left atrial volume and epicardial fat) were disregarded from evidence to calculate conditional probabilities.

The performance of the model for classifying AF relapse with these 5 predictors was as expectably lower than with 7 predictors, with a calculated AUC average of 0.661 (95% CI 0.603 - 0.718) for all sampling methods. ROC curves for each validation test are shown in Figure 4.

**Figure 4.** Receiver operating characteristic curves for all validation sampling methods applied to the model with 5 predictors: age, sex, smoking, preablation atrial fibrillation type, and obstructive sleep apnea. AUC values averaged 0.661 (95% CI 0.603 - 0.718). AUC: area under the curve.




## The 6 Predictors

The predictive performance with only the previous 5 predictors appears to be slightly more than average. However, it can be observed from the defined Bayesian network structure (Figure 1) that the epicardial fat node has BMI as its single parent, meaning that the latter directly influences the former. As such, the lack of information on epicardial fat for a given patient can be partially compensated by its information on the BMI value. This poses an interesting possibility, especially when observed that BMI is usually an available or easy to obtain feature for any patient.

The rationale for this test was therefore to gauge the predictive power of a model when using the 5 predictors in the previous experience, plus the information on the BMI node. All these 6 features—age, sex, smoking, preablation AF type, OSA, and BMI—are usually easily available clinical variables for physicians' evaluation, which do not require the use of additional complex or expensive diagnostic means. Therefore, this setting simulates the predictive power of the model in a likely real-life scenario.

For this test, we calculated the conditional probability of AF relapse for each patient in the test set, considering evidence on age, sex, smoking, preablation AF type, OSA, and BMI. Any information on left atrial volume and epicardial fat was ignored for this purpose.

The performance of the model for classifying AF relapse with these 6 predictors resulted in a computed AUC average of 0.703 (95% CI 0.652 - 0.753) for all sampling methods. ROC curves for each validation test are shown in Figure 5.

**Figure 5.** Receiver operating characteristic curves for all validation sampling methods applied to the model with 6 predictors: age, sex, smoking, preablation atrial fibrillation type, obstructive sleep apnea, and BMI. AUC values averaged 0.703 (95% CI 0.652 - 0.753). AUC: area under the curve.

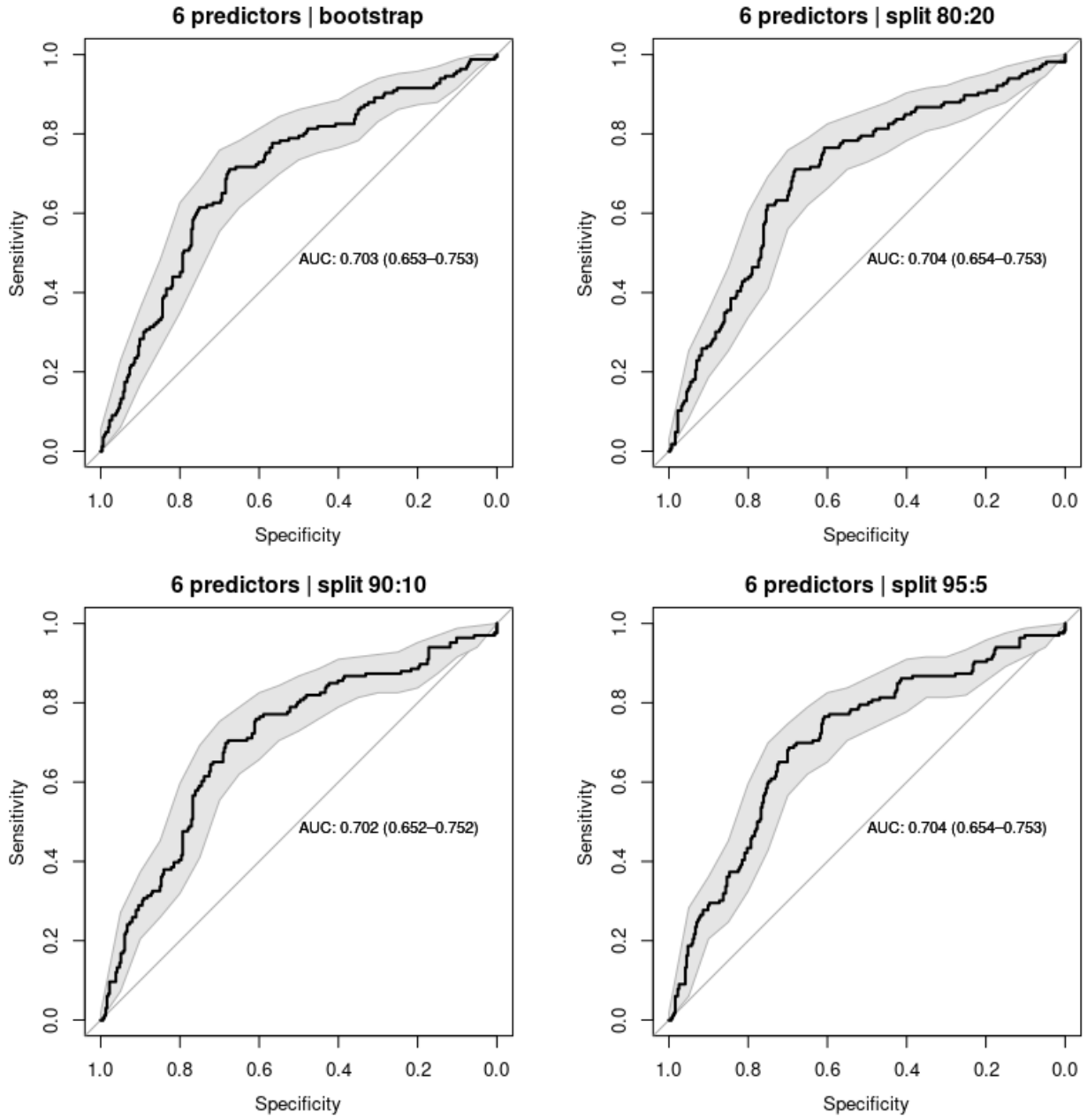



Table 3 presents a comparative analysis of the three models developed using 5, 6, and 7 predictors, respectively. As shown, the AUC-ROC progressively increases with the addition of predictors, indicating improved model performance. Furthermore, the 95% CI narrows as the number of predictors increases, suggesting greater precision in the model's estimates.

**Table .** Comparative analysis of model performance based on the number of predictors and validation sampling techniques, using area under the receiver operating characteristic curve (AUC-ROC) metrics.

| Model | AUC-ROC (95% CI) | | | | |
|---|---|---|---|---|---|
| | Bootstrap | Split 80:20 | Split 90:10 | Split 95:5 | Mean |
| 5 predictors | 0.658 (0.603 - 0.713) | 0.660 (0.605 - 0.715) | 0.660 (0.605 - 0.715) | 0.664 (0.610 - 0.718) | 0.661 (0.603 - 0.718) |
| 6 predictors | 0.703 (0.653 - 0.753) | 0.704 (0.654 - 0.753) | 0.702 (0.652 - 0.752) | 0.704 (0.654 - 0.753) | 0.703 (0.652 - 0.753) |
| 7 predictors | 0.755 (0.710 - 0.800) | 0.751 (0.706 - 0.797) | 0.747 (0.701 - 0.792) | 0.755 (0.709 - 0.800) | 0.752 (0.701 - 0.800) |

## Discussion

### Principal Findings

The ability to accurately predict clinical outcomes is vital for improving the quality of medical care and increasing the efficiency of resource allocation in health care. For such predictions, cardiologists often use clinical scores that have various limitations, such as being dependent on a set number of medical variables or not being adaptable to new medical knowledge. Nonetheless, these professionals have also been witnessing the development of AI models for applications in cardiology in general [20] and for the management of arrhythmias in particular [21,22]. In this context, our aim was to develop an alternative model to clinical scores that was not susceptible to these limitations, to predict the relapse of AF after PVI procedure.

For this purpose, we have resorted to Bayesian networks, a type of probabilistic graphical model that can represent knowledge as a set of variables and their conditional dependencies. Unlike traditional prognostic models based on linear or logistic regressions, Bayesian networks offer an interpretable graphical structure, which enhances the model's clarity and facilitates its adoption among physicians. In addition, Bayesian networks manage missing data more efficiently than other machine learning methods like classification and regression trees or random forests, as they can compute the probability of an outcome even when predictive variables have missing values. This makes them particularly well suited for medical datasets, where missing data are often a challenge. We have therefore chosen to develop our models based on Bayesian networks due to their explainability, flexibility, and robustness. Their explainability derives from their ability to represent relationships between variables as a graphical model, thus rendering their results more comprehensible. This capability is of paramount importance for the acceptance of AI models by medical professionals, who can thus integrate them safely into clinical practice [23]. Further, the models' flexibility derives from the ability to accommodate and represent new medical knowledge by reshaping the network structure accordingly and recalculating the conditional dependencies among multiple variables. Therefore, new suspected or known risk factors or predictors for AF relapse can be incorporated into a Bayesian network model at any time, with minimal resetting of the model. Additionally, the models' robustness derives from the fact that they can make predictions for the outcome variable even when there are missing data on some predictive variables, thus allowing them to be used in cases of incomplete information on any given patient. Thus, unlike clinical scores, Bayesian networks do not require the full set of clinical explanatory variables to deliver useful results. Despite none of these characteristics being unique to Bayesian networks on its own, this combination of characteristics makes these models highly interesting to be used as basis for clinical decision support tools.The first stage of the construction of our model was to create the network structure, that is, the network of relationships between the clinical variables. As stated in the *Methods* section, this was achieved in 2 steps: initially the known relationships were set manually based on expert knowledge; then, in a second step, the network structure was improved upon inference from data by the use of an AI algorithm. At this last step, the algorithm suggested a relationship between BMI and epicardial fat, which was considered acceptable, as there is significant evidence of a correlation between these two variables [24]. This finding proved useful since it enabled the use of the path "BMI → epicardial fat → AF relapse" when there was no information on the middle variable. The algorithm also suggested a path "OSA → pre-ablation AF type → left atrial volume." In this study, we opted to retain this suggestion in the network structure as a potential motivation for further exploration in future research. Although these relationships were considered to represent knowledge derived from the data, they were not particularly relevant for the model calculations, since each of these variables is also directly related to the outcome variable.

The second stage of the construction of our model was to train and validate the model based on the previous network structure. When validating the use of evidence from the 7 parent nodes of our outcome variable, the model performed with a calculated AUC value of approximately 0.75, interpreted as acceptable diagnostic accuracy [25]. These results implied using as predictive variables age, sex, smoking, preablation AF type, left atrial volume, epicardial fat, and OSA. However, some of these features are not always available in patients' clinical records. Thus, we have validated the model in the absence of information on left atrial volume and epicardial fat as predictive features. In this case, the model exhibited an expectedly lower performance, with a calculated mean AUC value close to 0.66. Despite the observed difference was not statistically significant, as noted from the overlapping confidence intervals, it suggests that these 2 features have a high weight on the performance of the model. This finding is consistent with those reported in the ATLAS score that the left atrial volume has the highest weight on the predictive power of that score [2].

Going further, our experiment also showed that the lack of information on epicardial fat can be partially compensated for by evidence of BMI, as this is its parent node. Taking into account daily clinical practice, this poses an interesting possibility, since BMI measurements are generally available for clinical evaluation for most patients. In these 6-variable cases, the model response exhibited a calculated mean AUC value of 0.70. Also here, despite the observed differences for the previous scenarios not being statistically significant, these outcomes fit within an acceptable range for a prediction tool. Such results implied using as predictive variables age, sex, smoking, preablation AF type, OSA, and BMI, all of which are typically easy to obtain in a clinical setting. To put these results in perspective, the AFA Recur tool developed by Saglietto et al [5] achieves a performance of AUC 0.72 using a 19-variable AI model with little to no explainability.

Future research in the context of predicting AF relapse using Bayesian networks should address several key challenges and directions. The first is ensuring the generalizability of the model across diverse populations and clinical settings to seek validation in varied patient cohorts. Second, it would be essential to conduct longitudinal studies to assess the model's long-term performance and capture patient evolution over extended time horizons. In addition, future studies could explore the inclusion

of expanded predictive factors, such as genetic influences, lifestyle changes, and comorbidities, to enhance the model's accuracy and clinical use. Finally, incorporating patient-reported outcomes and preferences into the predictive framework may improve the model's relevance and acceptance, fostering a more patient-centric approach to clinical decision-making.

We consider that this data-based approach based on a Bayesian network model can be the backbone for a future clinical decision support system. Being an AI model, it opens the possibility of being continuously retrained as new patient information becomes available in clinical records, hence progressively providing more accurate results upon new accumulated data. Such a retraining process can be automatized on a schedule or upon a trigger, for example, recalculating conditional dependencies between clinical features on a monthly basis or at every new 100 patient observations. This retraining of the model based on the recalculation of conditional probabilities from new patient data is not expected to represent significant computational costs, even for exceptionally large amounts of patient observations.

This model can also be considered as an enhancement of the ATLAS score, as it is based on its 5 predictive features, to which 2 additional features were added. Nonetheless, it may serve as a starting point for the representation of knowledge in this field, being open to incorporating new evidence as it becomes available. For such a reason, we believe that the findings of this research contribute to the growing body of knowledge on the application of AI methods in cardiology and pave the way for future advancements in predictive analytics for cardiovascular diseases.

### Strengths and Limitations

The model was developed and evaluated on a dataset with a limited number of features. Although the current literature identifies other potential risk factors for relapse of AF, these were not considered in this work, as there was no information from patients on such features. Nevertheless, this type of model allows the incorporation of other risk factors at any time, provided that the network structure is rebuilt for that knowledge representation and the model is retrained accordingly.

In addition, the size of the dataset used in this work was below optimal for this type of probabilistic model. This is particularly relevant if we consider the subsample sizes for a given combination of clinical conditions (eg, in this dataset, there was only one observation that simultaneously satisfies the multiple conditions sex = female + smoking = true + OSA = true). However, this type of model can be set to learn from new patient data as they becomes available. In this fashion, as it continuously builds on new evidence, the model becomes more accurate and reliable, even for less frequent clinical conditions.

## Acknowledgments

This work has been carried out within the scope of the PhD Programme in Health Data Science of the Faculty of Medicine of the University of Porto, Portugal.

## Conflicts of Interest

None declared.

## References

1. Rosenberg MA, Patton KK, Sotoodehnia N, et al. The impact of height on the risk of atrial fibrillation: the Cardiovascular Health Study. Eur Heart J 2012 Nov;33(21):2709-2717. [doi: 10.1093/eurheartj/ehs301] [Medline: 22977225]
2. Mesquita J, Ferreira AM, Cavaco D, et al. Development and validation of a risk score for predicting atrial fibrillation recurrence after a first catheter ablation procedure - ATLAS score. Europace 2018 Nov 1;20(FI_3):f428-f435. [doi: 10.1093/europace/eux265] [Medline: 29016770]
3. Gill SK, Karwath A, Uh HW, et al. Artificial intelligence to enhance clinical value across the spectrum of cardiovascular healthcare. Eur Heart J 2023 Mar 1;44(9):713-725. [doi: 10.1093/eurheartj/ehac758] [Medline: 36629285]
4. Tarakji KG, Silva J, Chen LY, et al. Digital health and the care of the patient with arrhythmia. Circ Arrhythm Electrophysiol 2020 Nov;13(11):13. [doi: 10.1161/CIRCEP.120.007953]
5. Saglietto A, Gaita F, Blomstrom-Lundqvist C, et al. AFA-Recur: an ESC EORP AFA-LT registry machine-learning web calculator predicting atrial fibrillation recurrence after ablation. EP Europace 2023 Feb 8;25(1):92-100. [doi: 10.1093/europace/euac145]
6. Budzianowski J, Kaczmarek-Majer K, Rzeźniczak J, et al. Machine learning model for predicting late recurrence of atrial fibrillation after catheter ablation. Sci Rep 2023 Sep 14;13(1):15213. [doi: 10.1038/s41598-023-42542-y] [Medline: 37709859]
7. Li G, Wang X, Han JJ, Guo X. Development and validation of a novel risk model for predicting atrial fibrillation recurrence risk among paroxysmal atrial fibrillation patients after the first catheter ablation. Front Cardiovasc Med 2022;9:1042573. [doi: 10.3389/fcvm.2022.1042573]
8. Ogbomo-Harmitt S, Muffoletto M, Zeidan A, Qureshi A, King A, Aslanidi O. Can artificial intelligence prediction of successful atrial fibrillation catheter ablation therapy be interpretable? Eur Heart J 2022 Oct 3;43(Supplement_2). [doi: 10.1093/eurheartj/ehac544.2775]

9. Pecha S. Can we predict the risk for atrial fibrillation recurrence after concomitant surgical atrial fibrillation ablation and should this influence our treatment strategy? Eur J Cardiothorac Surg 2021 Jun 14;59(6):1225-1226. [doi: 10.1093/ejcts/ezab022] [Medline: 34007985]
10. Kyrimi E, Dube K, Fenton N, et al. Bayesian networks in healthcare: what is preventing their adoption? Artif Intell Med 2021 Jun;116:102079. [doi: 10.1016/j.artmed.2021.102079] [Medline: 34020755]
11. Nascimento Matos D, Ferreira AM, Cavaco D, et al. Epicardial fat volume outperforms classic clinical scores for predicting atrial fibrillation relapse after pulmonary vein isolation. Eur Heart J 2020 Nov 1;41(Supplement_2). [doi: 10.1093/ehjci/ehaa946.0587]
12. Wong CX, Sun MT, Odutayo A, et al. Associations of epicardial, abdominal, and overall adiposity with atrial fibrillation. Circ Arrhythm Electrophysiol 2016 Dec;9(12):9. [doi: 10.1161/CIRCEP.116.004378] [Medline: 27923804]
13. Chen W, Cai X, Yan H, Pan Y. Causal effect of obstructive sleep apnea on atrial fibrillation: a Mendelian randomization study. J Am Heart Assoc 2021 Dec 7;10(23):e022560. [doi: 10.1161/JAHA.121.022560] [Medline: 34796736]
14. Linz D, Nattel S, Kalman JM, Sanders P. Sleep apnea and atrial fibrillation. Card Electrophysiol Clin 2021 Mar;13(1):87-94. [doi: 10.1016/j.ccep.2020.10.003] [Medline: 33516410]
15. Schwarz GE. Estimating the dimension of a model. Ann Statist 1978;6(2):461-464. [doi: 10.1214/aos/1176344136]
16. Russell SJ, Norvig P, Davis E. Artificial Intelligence: A Modern Approach, 3rd edition: Prentice Hall; 2010.
17. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2022. URL: https://cran.r-project.org/bin/windows/base/old/4.2.2/ [accessed 2025-01-27]
18. Scutari M. Learning Bayesian networks with the bnlearn R package. J Stat Softw 2010;35(3). [doi: 10.18637/jss.v035.i03]
19. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011 Mar 17;12(1):77. [doi: 10.1186/1471-2105-12-77] [Medline: 21414208]
20. Van den Eynde J, Lachmann M, Laugwitz KL, Manlhiot C, Kutty S. Successfully implemented artificial intelligence and machine learning applications in cardiology: state-of-the-art review. Trends Cardiovasc Med 2023 Jul;33(5):265-271. [doi: 10.1016/j.tcm.2022.01.010] [Medline: 35101642]
21. Nagarajan VD, Lee SL, Robertus JL, Nienaber CA, Trayanova NA, Ernst S. Artificial intelligence in the diagnosis and management of arrhythmias. Eur Heart J 2021 Oct 7;42(38):3904-3916. [doi: 10.1093/eurheartj/ehab544] [Medline: 34392353]
22. Feeny AK, Chung MK, Madabhushi A, et al. Artificial intelligence and machine learning in arrhythmias and cardiac electrophysiology. Circ Arrhythm Electrophysiol 2020 Aug;13(8):e007952. [doi: 10.1161/CIRCEP.119.007952] [Medline: 32628863]
23. Arbelaez Ossa L, Starke G, Lorenzini G, Vogt JE, Shaw DM, Elger BS. Re-focusing explainability in medicine. Dig Health 2022;8:20552076221074488. [doi: 10.1177/20552076221074488] [Medline: 35173981]
24. Shetty R, Vivek G, Naha K, Nayak K, Goyal A, Dias LS. Correlation of epicardial fat and anthropometric measurements in Asian-Indians: a community based study. Avicenna J Med 2012 Oct;2(4):89-93. [doi: 10.4103/2231-0770.110739] [Medline: 23826555]
25. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010 Sep;5(9):1315-1316. [doi: 10.1097/JTO.0b013e3181ec173d] [Medline: 20736804]


## Abbreviations

**AF:** atrial fibrillation
**AI:** artificial intelligence
**AUC:** area under the curve
**AUC-ROC:** area under the receiver operating characteristic curve
**BIC:** Bayesian Information Criterion
**CT:** computed tomography
**HBP:** high blood pressure
**OSA:** obstructive sleep apnea
**PV:** pulmonary vein
**PVI:** pulmonary vein isolation
**ROC:** receiver operating characteristic

*Edited by A Coristine; submitted 11.04.24; peer-reviewed by K Qu, K Chadaga; revised version received 19.11.24; accepted 19.11.24; published 11.02.25.*

*Please cite as:*
*Alves JM, Matos D, Martins T, Cavaco D, Carmo P, Galvão P, Costa FM, Morgado F, Ferreira AM, Freitas P, Dias CC, Rodrigues PP, Adragão P*
*Predicting Atrial Fibrillation Relapse Using Bayesian Networks: Explainable AI Approach*
*JMIR Cardio 2025;9:e59380*
*URL: https://cardio.jmir.org/2025/1/e59380*
*doi:10.2196/59380*

# Photoplethysmography-Based Machine Learning Approaches for Atrial Fibrillation Burden: Algorithm Development and Validation

Hong Wang[1,2], MM; Binbin Liu[1,2], MM; Hui Zhang[1], MD; Zheqi Zhang[1,2], MM; Zhigeng Jin[1], MD; Hao Wang[3], Prof Dr; Yu-Tao Guo[1], Prof Dr

[1]Department of Pulmonary Vessel and Thrombotic Disease, Sixth Medical Center, Chinese PLA General Hospital, 6 Fucheng Road, Haidian District, Beijing, China

[2]Graduate School of PLA General Hospital, Beijing, China

[3]Department of Cardiology, Second Medical Centre, National Clinical Research Centre for Geriatric Diseases, Chinese PLA General Hospital, Beijing, China

**Corresponding Author:**
Yu-Tao Guo, Prof Dr
Department of Pulmonary Vessel and Thrombotic Disease, Sixth Medical Center, Chinese PLA General Hospital, 6 Fucheng Road, Haidian District, Beijing, China

## Abstract

**Background:** Atrial fibrillation (AF) burden is associated with cardiovascular events such as stroke and heart failure. Recent advancements in photoplethysmography (PPG) technology have provided new insights into noninvasive and convenient AF burden detection.

**Objective:** This study aimed to establish an AF burden model based on smartwatch-monitored PPG technology to track the progression of AF.

**Methods:** This prospective pilot study (January 2024 to January 2025) at the Chinese PLA General Hospital enrolled patients with paroxysmal AF. Participants underwent simultaneous rhythm monitoring using smartwatch PPG and 24-hour Holter electrocardiogram monitoring (the gold standard). Five PPG-derived AF burden metrics were defined: (1) ratio of AF episode duration to total monitoring time (M1), (2) ratio of AF episode frequency to total measurements (M2), (3) AF episode density (M3), (4) AF episode variability (M4), and (5) proportion of rapid ventricular rate in AF episodes (>120 beats per minute; M5). Smartwatch PPG signals were collected once per minute. Sensitivity, specificity, accuracy, precision, and $F_1$ score were used to evaluate the PPG algorithm's AF detection capability through comparison with the gold standard (24-hour Holter monitoring). The mean absolute error (MAE) and Spearman rank correlation coefficient ($r_s$) were used to assess the correlation between the PPG-based AF burden metrics and the gold standard.

**Results:** A total of 145 participants with paroxysmal AF (n=96, 66.2% male; mean age 63.28, SD 14.23 years) were included. Compared to the gold standard, the PPG-based AF burden model demonstrated a sensitivity of 91.5% (95% CI 87.9%-95.1%), specificity of 97.2% (95% CI 95.9%-98.5%), precision of 92.9% (95% CI 88.6%-97.3%), accuracy of 93.3% (95% CI 88.2%-98.5%), and $F_1$ score of 90.5% (95% CI 86.3%-94.7%). The AF burden model exhibited strong discriminatory power in the test cohort (area under the curve=89.5%, 95% CI 89.4% - 89.7%). For M1, the MAE for the model of AF episode duration as a proportion of total monitoring time was 0.0400 ($P$=.008), with a correlation coefficient ($r_s$) of 0.8788 ($P$<.001). For M4, the MAE for the AF episode variability model was 3.9967 ($P$<.001), with a correlation coefficient ($r_s$) of 0.7876 ($P$<.001). The MAE for the average real variability model was 4.6436 ($P$<.001), with a correlation coefficient ($r_s$) of 0.8127 ($P$<.001). The MAE for the average AF change model was 0.3893 ($P$=.27), with a correlation coefficient ($r_s$) of 0.7246 ($P$<.001).

**Conclusions:** The PPG-based AF burden model demonstrated high concordance with the gold standard of 24-hour Holter monitoring in tracking AF episode duration and variability, providing new perspectives for exploring AF progression dynamics.

**Trial Registration:** Chinese Clinical Trial Registry ChiCTR2300075516; https://www.chictr.org.cn/showproj.html?proj=200976

***(JMIR Cardio 2025;9:e78075)*** doi:10.2196/78075

**KEYWORDS**

atrial fibrillation; photoplethysmography; atrial fibrillation burden; wearable devices; arrhythmia

## Introduction

Atrial fibrillation (AF) is one of the most common arrhythmias, associated with an increased risk of cardiovascular adverse events such as stroke and heart failure, and causes significant health, social, and economic burdens [1]. Numerous studies have found a dose-response relationship between AF burden and cardiovascular event risk, making early assessment of AF burden critical to reduce AF complications [2-4]. However, there is currently no standardized definition for AF burden. It generally refers to the percentage of time spent in AF during total monitoring duration. Other studies have alternatively quantified the number of AF episodes by using episode density to define AF burden to explore the relationship between AF burden and cardiovascular adverse events [2,5].

Currently, devices for monitoring AF burden include cardiovascular implantable electronic devices and noninvasive rhythm recorders (such as 24-hour Holter electrocardiograms [ECGs]), both of which demonstrate high diagnostic accuracy. However, these devices face limitations, including requiring hospital-based monitoring under clinician supervision, high costs, and labor-intensive data interpretation by physicians. Recently, photoplethysmography (PPG) has been developed for AF screening and has shown promising accuracy [6,7]. PPG devices for AF screening are commercially available in various formats, including handheld devices, smartwatches, or wristbands. They use built-in optical sensors to monitor blood volume changes in skin capillary beds and estimate heart rhythm via reflected light wavelengths, offering greater comfort and convenience and providing potential opportunities for out-of-hospital ambulatory AF burden assessment [8]. However, the accuracy of PPG-based AF burden detection algorithms in smartwatches remains uncertain [9].

Our study aimed to evaluate the accuracy of smartwatch integrated PPG algorithms in assessing AF burden.

## Methods

### Study Population

Between January 1, 2024, and January 1, 2025, consecutive patients diagnosed with paroxysmal AF were recruited from the Chinese PLA General Hospital. Inclusion criteria were patients aged ≥18 years who provided written informed consent. Exclusion criteria were inability to use wearable devices, cognitive impairment, or presence of implanted cardiac devices (pacemakers or implantable cardioverter-defibrillators).

### Ethical Considerations

This study complied with the World Medical Association's Declaration of Helsinki and was approved by the Institutional Review Board of the Chinese People's Liberation Army General Hospital (HZKY-PJ-2023-23). It was also registered with the Chinese Clinical Trial Registry (ChiCTR2300075516). All participants signed the informed consent form before participating in this study. This study strictly adhered to privacy protection protocols in accordance with the Declaration of Helsinki. No financial compensation was provided to participants.

### Signal Acquisition and Processing

This study involved collecting baseline clinical data and PPG signals. Clinical data included demographics, comorbidities, and medications. PPG signals were obtained as follows: after attaching 24-hour Holter ECG electrodes, participants wore smartwatches (Huawei Watch GT2 Pro; Huawei Technologies Co., Ltd.). Simultaneous recordings of cardiac rhythm from both devices were initiated.

### Development and Optimization of the Primary PPG-Based AF Burden Model

The PPG-based AF burden algorithm was developed using 3698 PPG data segments from the previous Mobile Atrial Fibrillation Application (mAFA) study as the training and validation sets [10]. Each data segment had a duration of 1 minute. The classification of the training data is shown in Table 1. The training data segments were each divided into varying durations (8, 16, 24, 32, 40, and 48 seconds) with random start times determined by a random number generator. During model development, we maintained a balanced representation of AF and non-AF episodes across all duration categories, enforced strict subject-level segregation to prevent any overlap between the training and validation datasets, and carefully preserved comparable AF to non-AF ratios in both sets. The raw PPG signal from the sensor module was processed using a bandpass Butterworth digital filter to eliminate low-frequency baseline drift and high-frequency noise, thereby obtaining clear and effective pulsatile PPG waveforms. To overcome limitations of single-time-AF detection algorithms, which exhibit low motion tolerance due to strict signal quality requirements and are inadequate for AF burden assessment, this study proposes a multiscale fusion AF burden detection algorithm featuring continuous PPG acquisition with minute-by-minute interpretation (defining AF burden as >40% of AF beats per minute), high-quality signal extraction from motion corrupted segments, an adaptive length machine learning model for variable duration PPG signals, and context-aware fusion calibration leveraging AF episode continuity to refine single time predictions, thereby achieving accurate AF burden quantification. In this study, we enrolled 145 patients as a test cohort to validate the detection accuracy of the AF burden algorithm compared to 24-hour Holter monitoring.

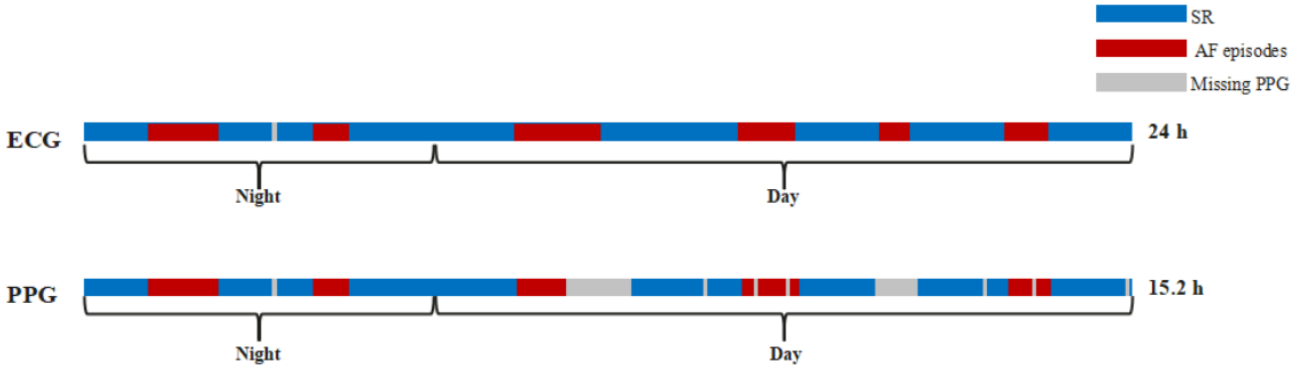
**Table .** The classification of the training data.

| Training data category | Record (1-minute segment) |
|---|---|
| Atrial fibrillation | 1090 |
| Premature contraction | 540 |
| Sinus rhythm | 2067 |

### Definition of AF Burden

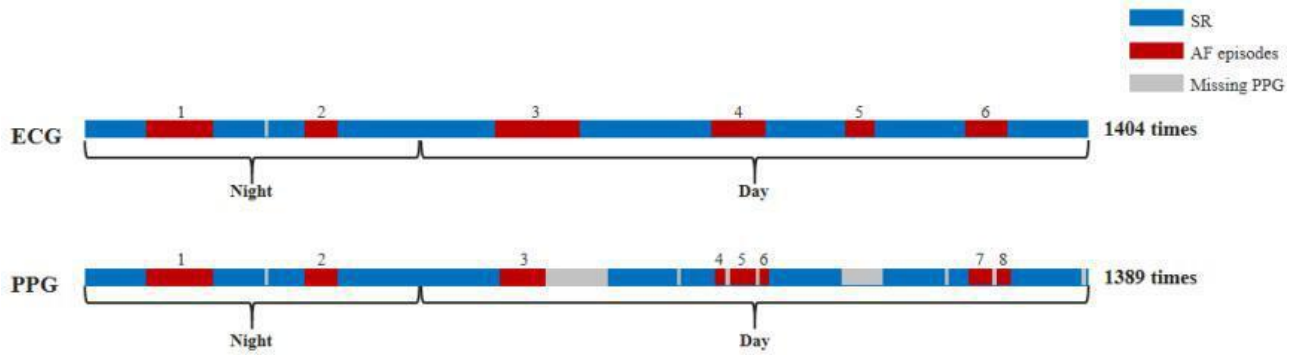
Metric M1 is the proportion of AF episode duration detected through PPG monitoring relative to the total monitoring time, quantifying temporal AF dynamics (Figure 1).

**Figure 1.** A schematic diagram of metric M1. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.



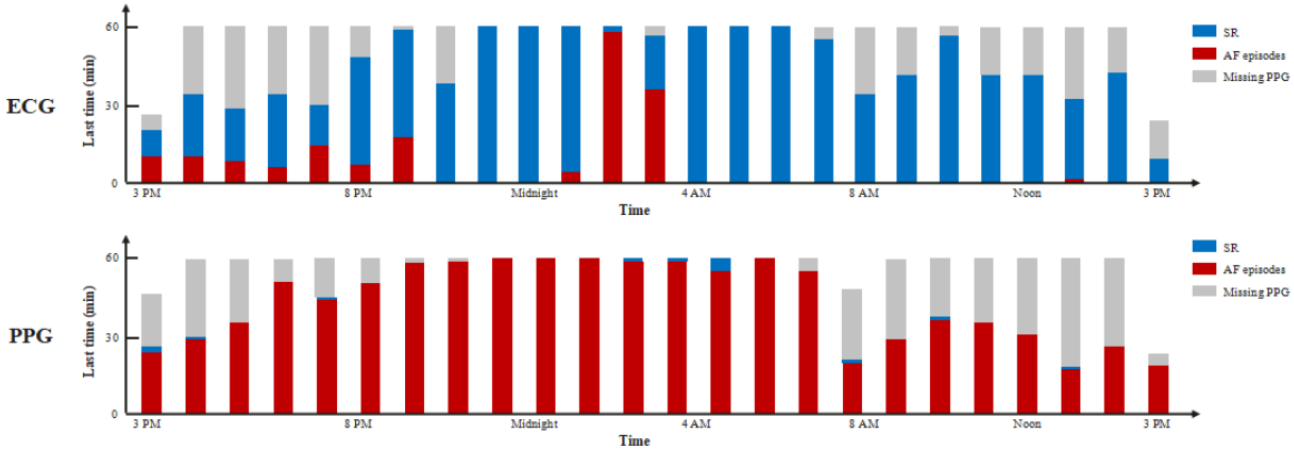

Metric M2 is the proportion of AF episodes detected through PPG monitoring relative to the total number of monitoring epochs, assessing frequency-based AF variations (Figure 2).

**Figure 2.** A schematic diagram of metric M2. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.



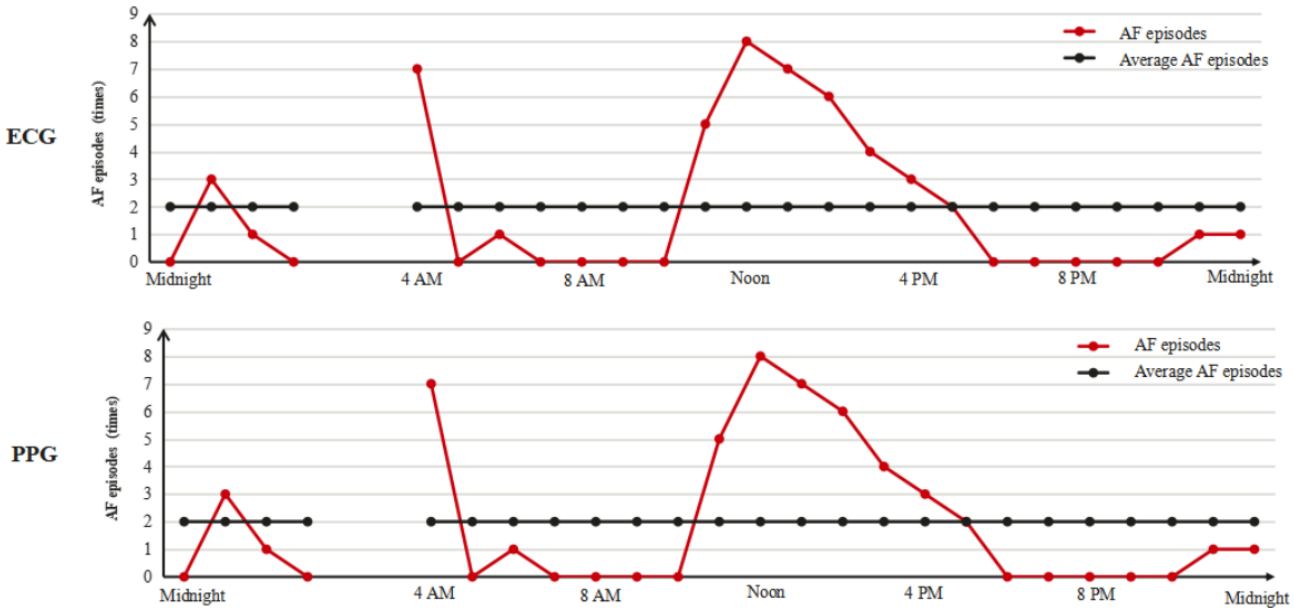

Metric M3 is the area between the actual AF burden progression curve (derived from PPG monitoring) and the theoretical uniform AF burden development curve, evaluating AF episode clustering (Figure 3).

**Figure 3.** A schematic diagram of metric M3. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.




Metric M4 is a composite metric integrating 3 variability parameters derived from PPG monitoring: AF episode variability (SD of hourly AF counts normalized by 24-hour mean AF frequency), mean real variability (average of SD and mean absolute deviation of hourly AF counts), and mean AF variation (day-night difference in AF counts normalized by overall mean AF frequency). This model quantifies hourly AF burden fluctuations (Figure 4).

**Figure 4.** A schematic diagram of metric M4. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.




Metric M5 quantifies rapid ventricular rate (RVR) AF episodes, defined as the proportion of AF episodes with heart rates of >120 beats per minute (bpm) relative to total monitored AF episodes or the cumulative duration of RVR AF episodes relative to the total monitoring time. This dual parameter model enables characterizing high rate AF burden (Figure 5 and Figure 6).

**Figure 5.** A schematic diagram of metric M5. AF: atrial fibrillation; bpm: beats per minute; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.

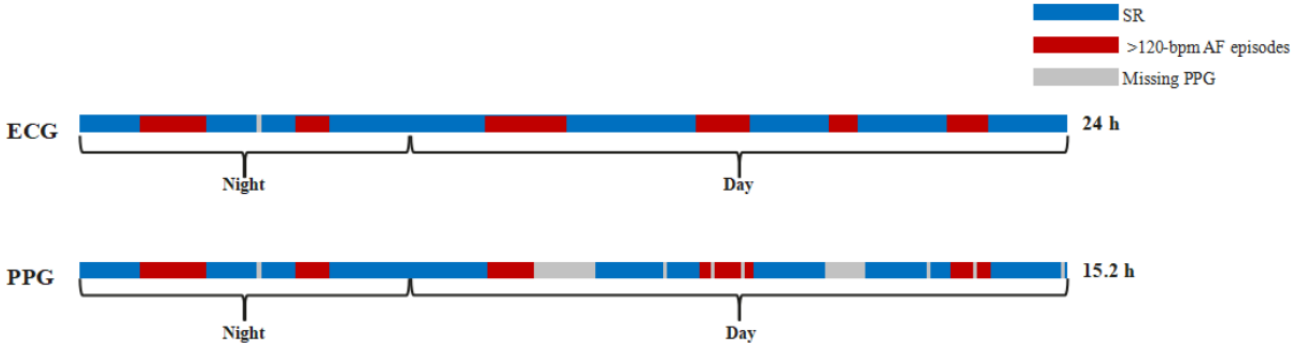
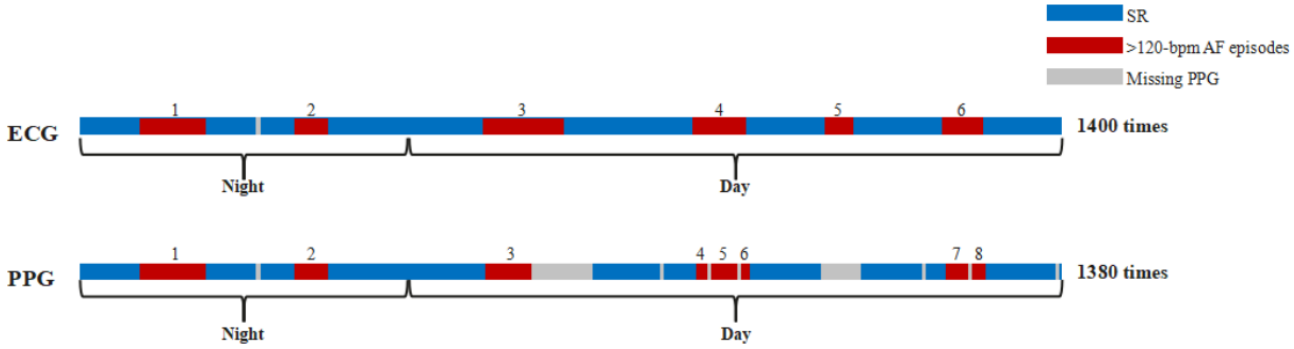



**Figure 6.** A schematic diagram of metric M5. AF: atrial fibrillation; bpm: beats per minute; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.




## Statistical Analysis

Data with a normal distribution were presented as means and SDs. Data with a nonnormal distribution were presented as medians and IQRs and were analyzed using the Mann-Whitney *U* test. Categorical variables were expressed as percentages. Algorithm performance (sensitivity, specificity, accuracy, precision, $F_1$ score, and area under the receiver operating characteristic curve) was validated against Holter ECG values to assess its capability for continuous AF detection. Correlation between the PPG-based AF burden model and 24-hour Holter ECG monitoring was assessed using the mean absolute error (MAE) and Spearman rank correlation coefficient ($r_s$; Table 2)

**Table .** Clinical significance of each metric.

| Performance metric | Formula | Clinical significance |
|---|---|---|
| Sensitivity | TP[a]/(TP + FN[b]) | Proportion of actual positive cases correctly identified |
| Specificity | TN[c]/(TN + FP[d]) | Proportion of actual negative cases correctly excluded |
| Accuracy | (TP + TN)/(TP + FP + FN + TN) | Proportion of correctly classified cases |
| Precision | TP/(TP + FP) | Proportion of TPs among all positive test results |
| $F_1$-score | 2 × TP/(2 × TP + FP + FN) | Balances FNs and FPs when their clinical consequences are comparable |
| AUC-ROC[e] | —[f] | Overall discriminative ability across all thresholds (0.9=excellent; 0.7 - 0.9=moderate) |
| MAE[g] | ∑i=1n\|pi−gi\|n,pi:PPGgi:ECG | Quantifies the average prediction error; smaller MAE indicates higher clinical utility for risk stratification or treatment timing |
| $r_s$ [h] | —f | Measures monotonic association; high $r_s$ ($r_s$>0.6) suggests that the model captures clinically relevant trends, even if they are nonlinear |

[a]TP: true positive.

[b]FN: false negative.

[c]TN: true negative.

[d]FP: false positive.

[e]AUC-ROC: area under the receiver operating characteristic curve.

[f]Not applicable.

[g]MAE: mean absolute error.

[h]$r_s$: Spearman rank correlation coefficient.

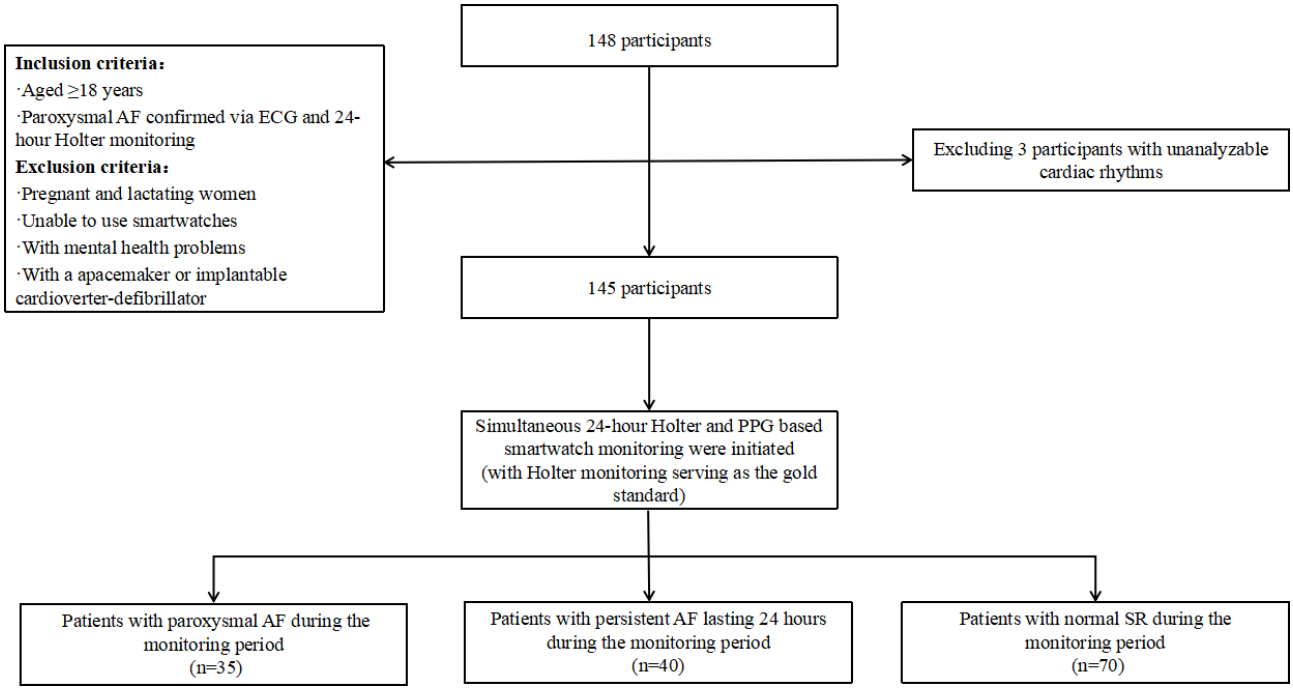
A 2-sided *P* value of <.05 was considered statistically significant. The 95% CIs were calculated using the Wilson score method without continuity correction. Analyses were conducted using SPSS Statistics (version 22; IBM Corp) and OpenEpi (version 3.01).

## Results

### Overview

From January 1, 2024, to January 1, 2025, a total of 148 participants were initially enrolled in the study. Of these 148 participants, after excluding 2 (1.4%) with atrial flutter and 1 (0.7%) with atrioventricular block, 145 (98%) were ultimately included in the final analysis (n=96, 66.2% male; mean age 63.28, SD 14.23; range 19 to 90 years). On the basis of expert-reviewed 24-hour Holter ECG monitoring, 75 patients exhibited AF episodes during the monitoring period, including 35 (47%) cases of paroxysmal AF and 40 (53%) cases of persistent AF lasting 24 hours (Figure 7).

**Figure 7.** A flowchart of the study. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.




## Baseline Characteristics

The study population comprised patients with paroxysmal AF with a mean age of 63.28 (SD 14.23) years, including 66.2% (96/145) male individuals. Baseline characteristics are detailed in Table 3.

**Table .** Baseline characteristics of the participants (N=145).

| Characteristic | Values |
|---|---|
| Demographics | |
| Age (y), mean (SD) | 63.28 (14.23) |
| Male sex, n (%) | 96 (66.2) |
| Medical history, n (%) | |
| Coronary artery disease | 31 (21.4) |
| Heart failure | 4 (2.8) |
| Hypertension | 75 (51.7) |
| Hyperlipidemia | 66 (45.5) |
| Diabetes mellitus | 31 (21.4) |
| Previous stroke, SE[a], or TIA[b] | 3 (2.1) |
| Vascular disease | 47 (32.4) |
| Renal dysfunction | 7 (4.8) |
| COPD[c] | 9 (6.2) |
| Hyperthyroidism | 3 (2.1) |
| OSA[d] | 15 (10.3) |
| Anticoagulants, n (%) | |
| Warfarin | 2 (1.4) |
| Dabigatran | 14 (9.7) |
| Rivaroxaban | 21 (14.5) |
| Apixaban | 3 (2.1) |
| Edoxaban | 24 (16.6) |
| Low–molecular weight heparin | 3 (2.1) |
| Antiarrhythmic drugs, n (%) | |
| Propafenone | 28 (19.3) |
| Amiodarone | 16 (11.0) |
| Dronedarone | 7 (4.8) |
| Bisoprolol | 17 (11.7) |
| Metoprolol | 39 (26.9) |

[a]SE: systemic arterial embolism.

[b]TIA: transient ischemic attack.

[c]COPD: chronic obstructive pulmonary disease.

[d]OSA: obstructive sleep apnea.

The AF burden algorithm demonstrated high performance, with a daily model output rate of 86.6% (95% CI 85.2%-88%). The confusion matrix yielded true positive, false negative, false positive, and true negative counts for PPG signals of 106,975, 11,941, 8787, and 308,082, respectively. Compared to the gold standard, the PPG-based AF burden model demonstrated a sensitivity of 91.5% (95% CI 87.9%-95.1%), specificity of 97.2% (95% CI 95.9%-98.5%), precision of 92.9% (95% CI 88.6%-97.3%), accuracy of 93.3% (95% CI 88.2%-98.5%), and $F_1$-score of 90.5% (95% CI 86.3%-94.7%). The AF burden model exhibited strong discriminatory power in the test cohort (area under the curve=89.5%, 95% CI 89.4%-89.7%; Table 4). The receiver operating characteristic curve is shown in Figure 8. The performance metrics of the PPG-based AF burden model across varying threshold values are shown in Table 5.

Figure 9 shows the strong concordance between the PPG and Holter ECG waveforms.

**Table .** The performance metrics of the photoplethysmography-based atrial fibrillation burden model.

| Performance metric | Performance (%; 95% CI) |
|---|---|
| Sensitivity | 90.0 (87.9 - 95.1) |
| Specificity | 97.2 (95.9 - 98.5) |
| Precision | 92.9 (88.6 - 97.3) |
| Accuracy | 93.3 (88.2 - 98.5) |
| $F_1$-score | 90.5 (86.3 - 94.7) |
| AUC[a] | 89.5 (89.4 - 89.7) |

[a]AUC: area under the receiver operating characteristic curve.

**Figure 8.** Receiver operating characteristic curves for photoplethysmography–based atrial fibrillation burden model. AUC: area under the curve; ROC: receiver operating characteristic.

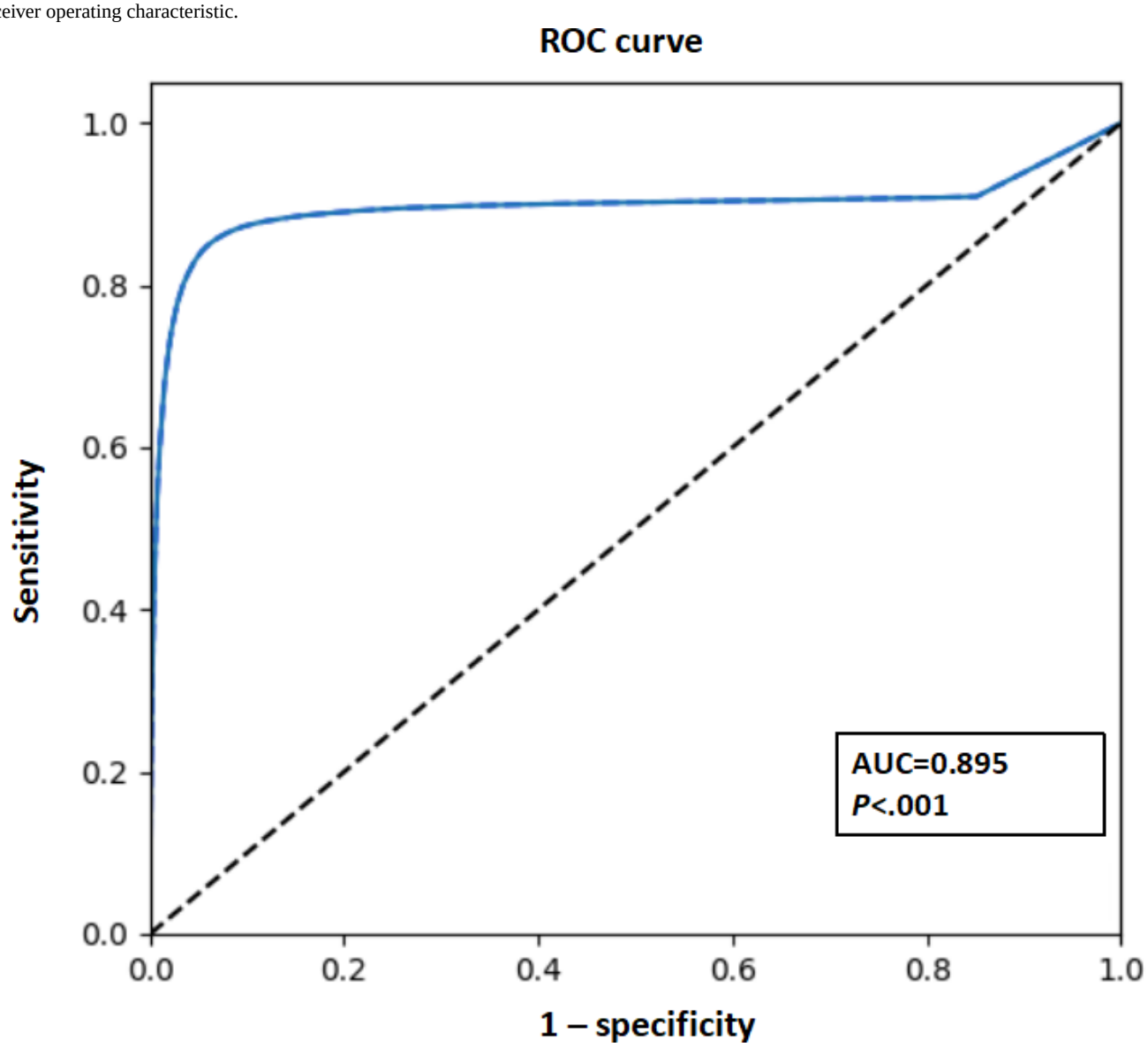
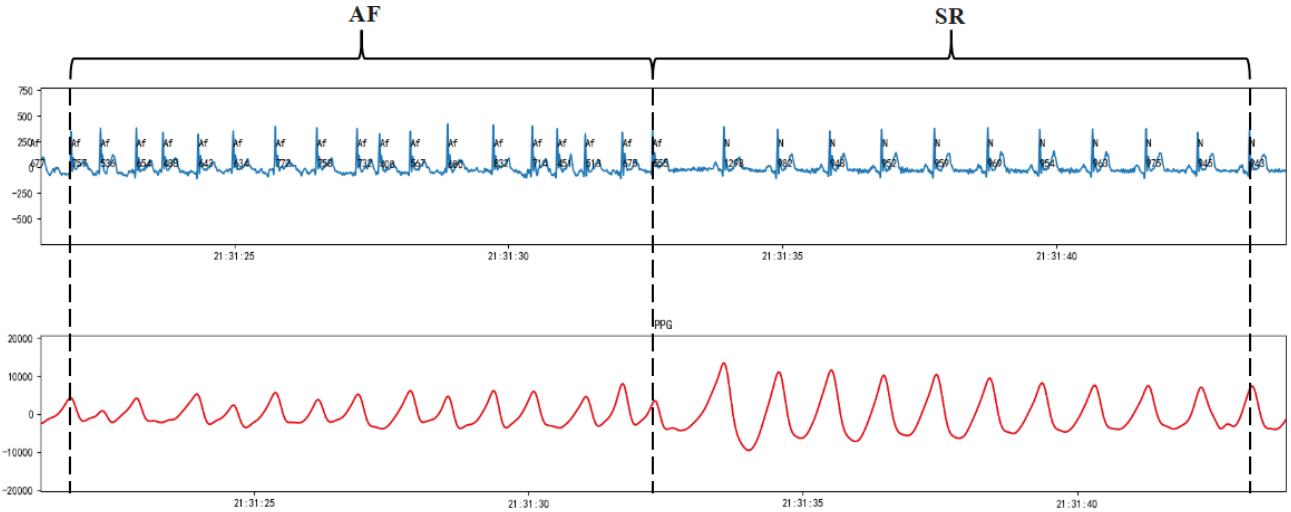


**Table .** The performance metrics of the photoplethysmography-based atrial fibrillation burden model across varying threshold values.

| Cutoff point for AF burden threshold | Accuracy | Precision | Sensitivity | Specificity | $F_1$ score |
|---|---|---|---|---|---|
| 40% | 0.91 | 0.85 | 0.84 | 0.94 | 0.84 |
| 45% | 0.92 | 0.85 | 0.84 | 0.94 | 0.85 |
| 50% | 0.92 | 0.84 | 0.85 | 0.94 | 0.85 |
| 55% | 0.92 | 0.84 | 0.86 | 0.94 | 0.85 |
| 60% | 0.92 | 0.83 | 0.87 | 0.94 | 0.85 |
| 65% | 0.92 | 0.83 | 0.87 | 0.94 | 0.85 |
| 70% | 0.92 | 0.82 | 0.88 | 0.93 | 0.85 |

**Figure 9.** Representative pulse waveform recording from a patient. AF: atrial fibrillation; PPG: photoplethysmography; SR: sinus rhythm.

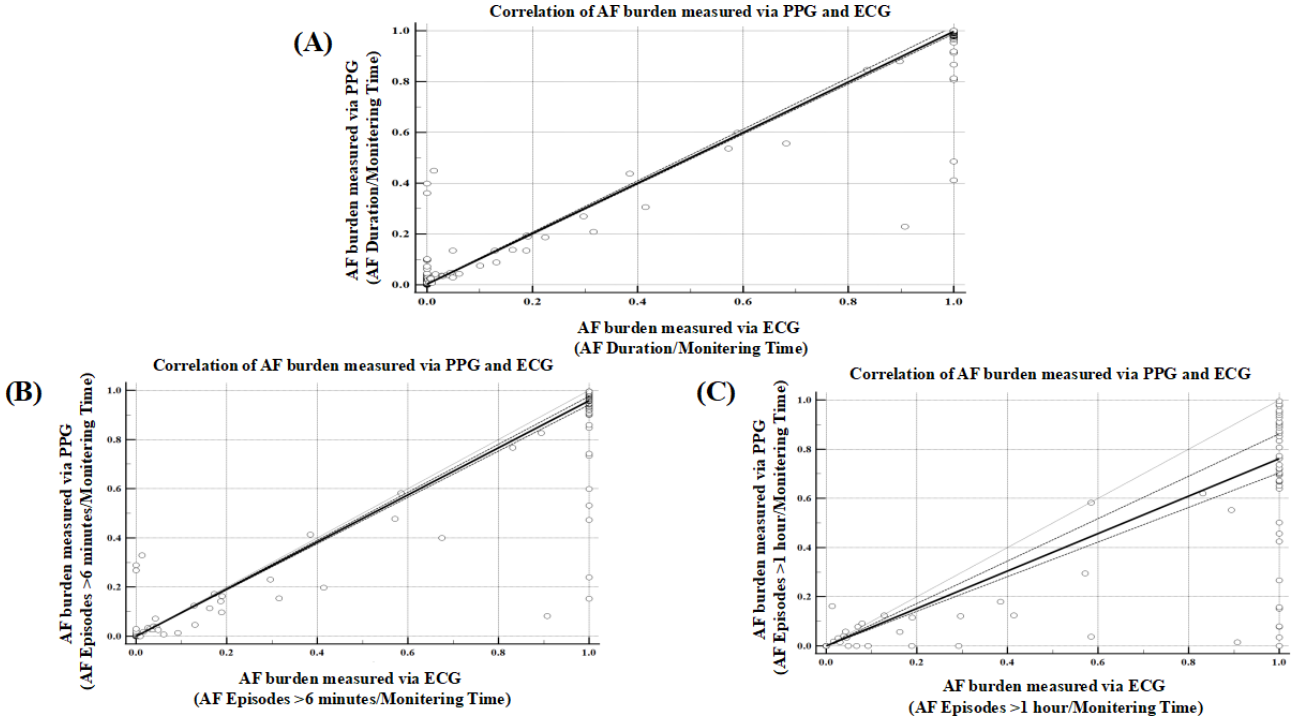



## Evaluation of AF Burden Quantification Performance by the AF Burden Model

### *Metric M1 Evaluation*

For the ratio of AF duration to the total monitoring time, the PPG median was 0.0402 (IQR 0.0086 - 0.8671) versus 0.0020 (IQR 0.0086 - 0.8671) for 24-hour Holter monitoring, with model performance metrics of MAE=0.0400 ($P$=.008) and $r_s$=0.8788 ($P$<.001; Figure 10A). For the ratio of AF episodes of >6 minutes to the total monitoring time, the PPG median was 0.0093 (IQR 0.0000 - 0.7330) versus 0.0000 (IQR 0.0000 - 1.0000) for 24-hour Holter monitoring, with model performance metrics of MAE=0.0582 ($P$=.44) and $r_s$=0.9233 ($P$<.001; Figure 10B). Regarding the ratio of AF episodes of >1 hour to the total monitoring time, the PPG median was 0.0000 (IQR 0.0000 - 0.2947) versus 0.0000 (IQR 0.0000 - 1.0000) for 24-hour Holter monitoring, with model performance of MAE=0.1204 ($P$=.04) and $r_s$=0.9293 ($P$<.001; Figure 10C).

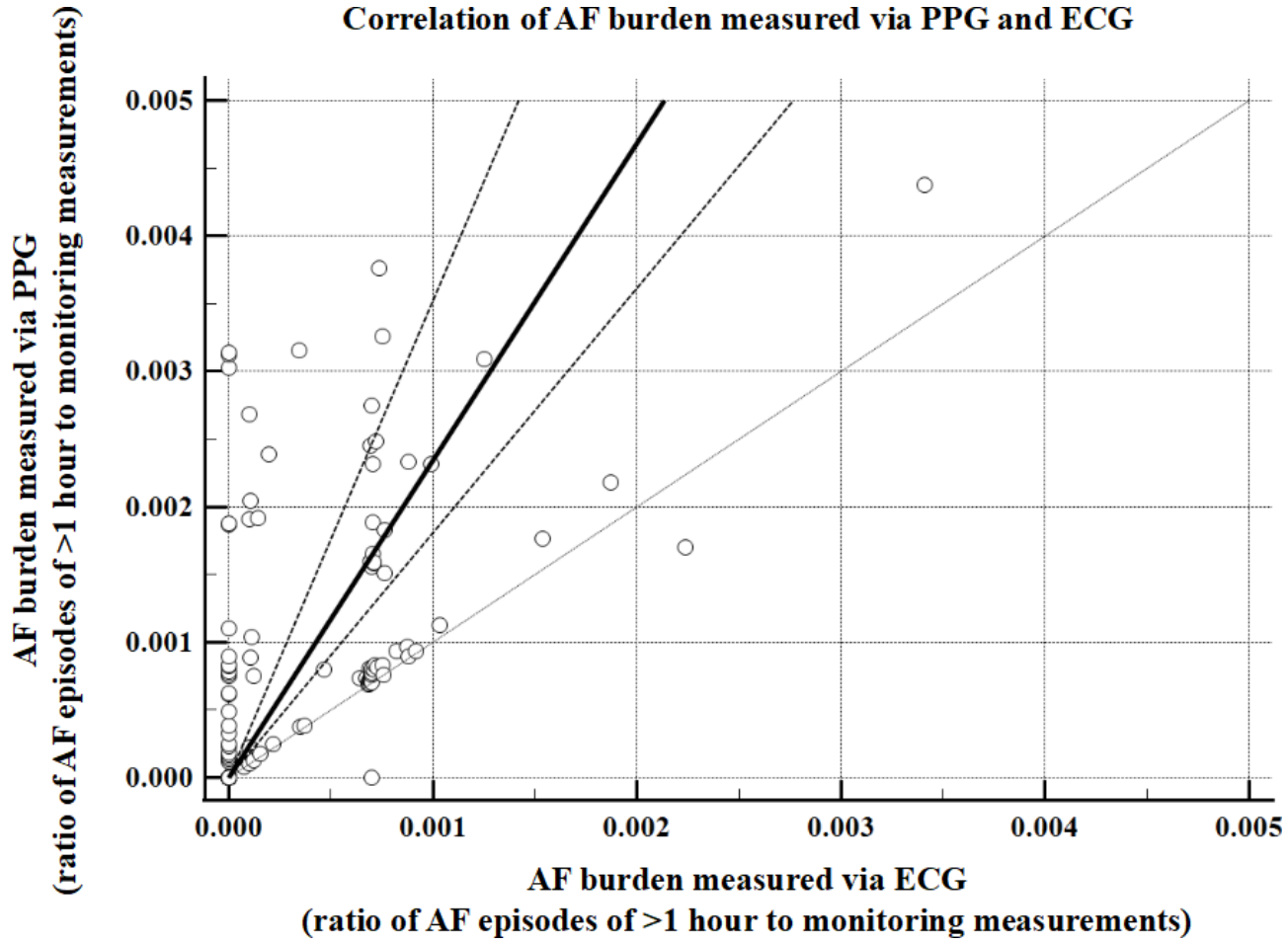
**Figure 10.** Scatter plot and regression line of AF burden estimated by PPG compared to AF burden estimated by ECG. Each dot represents the AF burden by each method for each patient with AF. Dashed line is the regression line. Solid gray line is line of equality. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography.




### *Metric M2 Evaluation*

The ratio of AF episodes to the total measurements showed a PPG median of 0.0033 (IQR 0.0009 - 0.0056) versus 0.0001 (IQR 0.0000 - 0.0008) for 24-hour Holter monitoring, with model performance metrics of MAE=0.0320 ($P$<.001) and $r_s$=–0.0807 ($P$=.33) compared to the gold standard. For the ratio of AF episodes of >6 minutes to the total measurements, the PPG median was 0.0010 (IQR 0.0007 - 0.0028) versus 0.0000 (IQR 0.0000 - 0.0007) for 24-hour Holter monitoring, with model performance metrics of MAE=0.0015 ($P$<.001) and $r_s$=0.2360 ($P$=.004). Regarding the ratio of AF episodes of >1 hour to the total measurements, the PPG median was 0.0004 (IQR 0.0000 - 0.0009) versus 0.0000 (IQR 0.0000 - 0.0007) for 24-hour Holter monitoring, with model performance of MAE=0.0005 ($P$<.001) and $r_s$=0.7092 ($P$<.001; Figure 11).

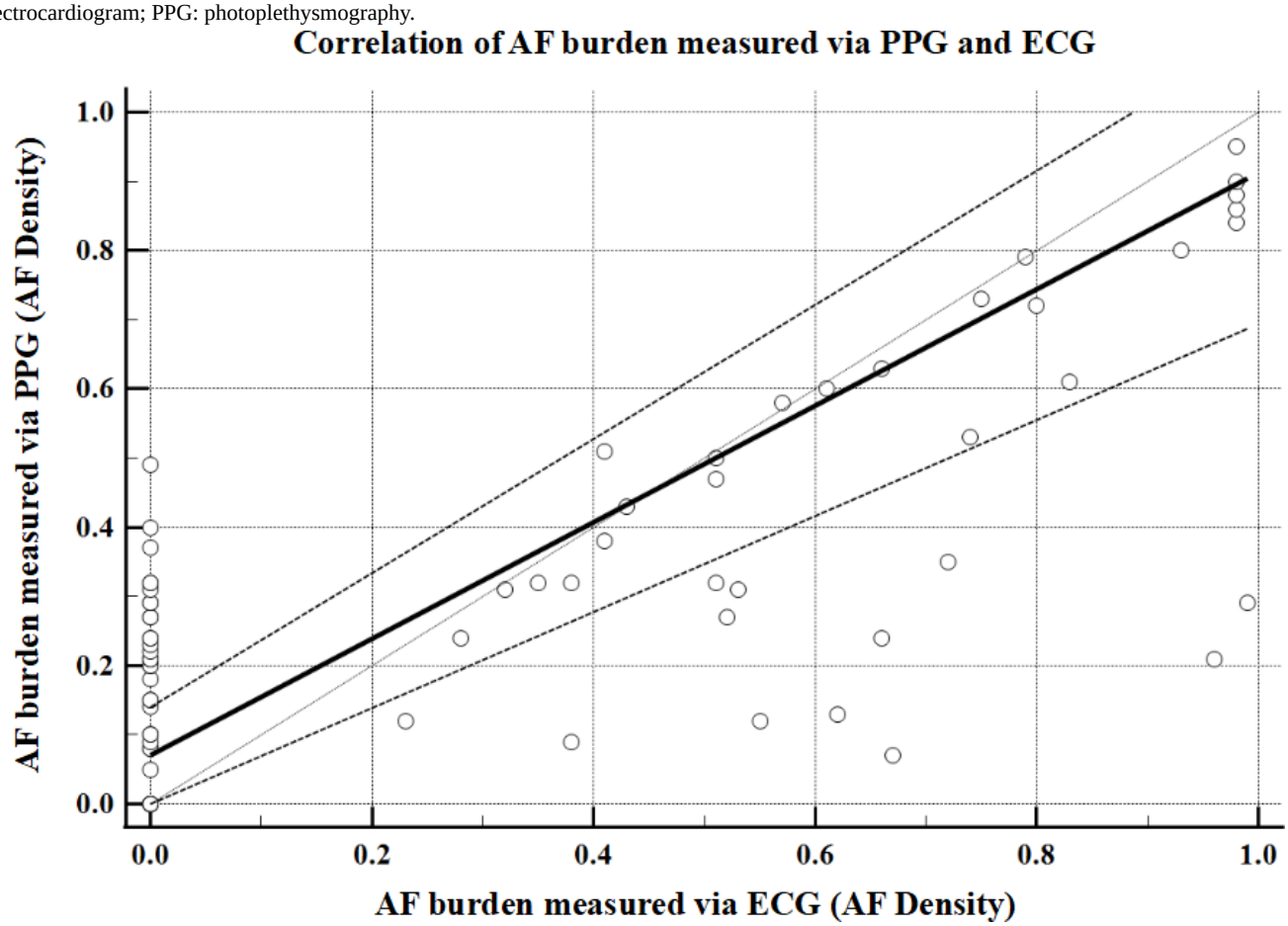
**Figure 11.** Scatter plot and regression line of AF burden estimated by PPG compared to AF burden estimated by ECG. Each dot represents the AF burden by each method for each patient with AF. Dashed line is the regression line. Solid gray line is line of equality. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography.




### *Metric M3 Evaluation*

The PPG-derived AF density showed a median of 0.2700 (IQR 0.1700-0.4725) compared to 0.0000 (IQR 0.0000-0.5900) for 24-hour Holter monitoring. When validated against the gold standard, the model demonstrated an MAE of 0.1725 ($P$<.001) with an $r_s$ of 0.6576 ($P$<.001; Figure 12)

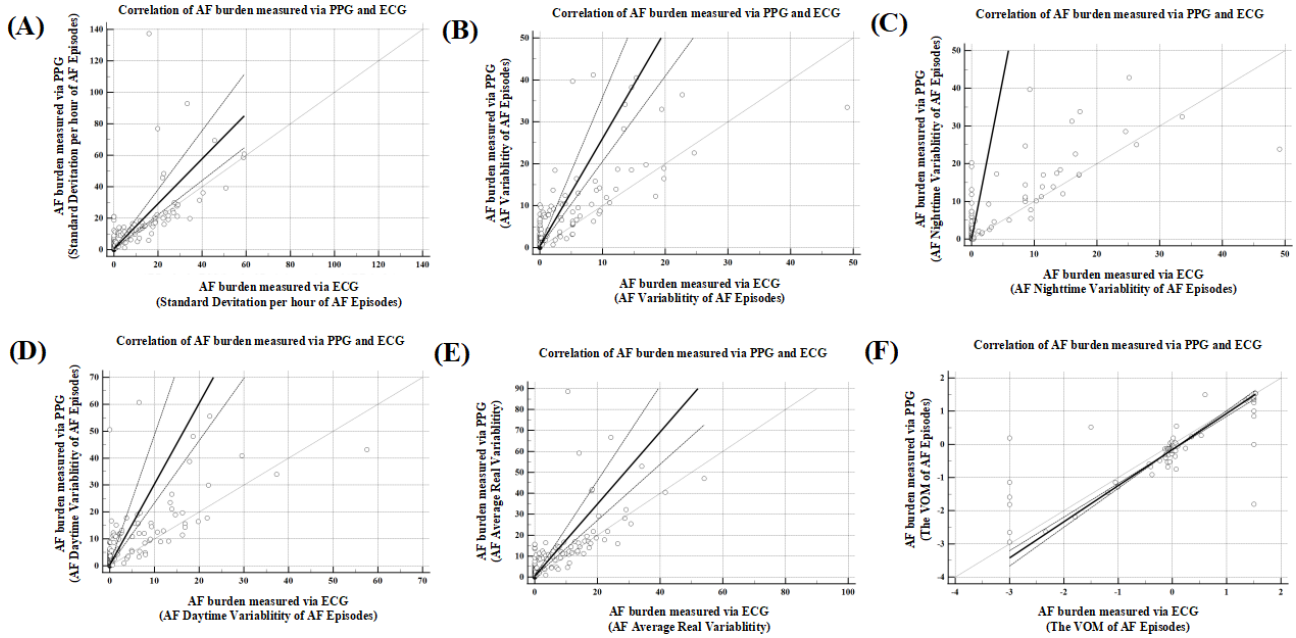
**Figure 12.** Scatter plot and regression line of AF burden estimated by PPG compared to AF burden estimated by ECG. Each dot represents the AF burden by each method for each patient with AF. Dashed line is the regression line. Solid gray line is line of equality. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography.




### *Metric M4 Evaluation*

The SD of hourly AF episodes measured through PPG showed a median of 5.3229 (IQR 1.0375-15.9058) versus 0.4390 (IQR 0.0000-12.8558) for Holter monitoring, with MAE=5.6009 ($P$<.001) and $r_s$=0.8135 ($P$<.001; Figure 13A). AF variability assessment revealed a median PPG value of 3.3478 (IQR 0.6957-9.2174) compared to 0.1739 (IQR 0.0000-5.2174) for 24-hour Holter monitoring, with MAE=3.9967 ($P$<.001) and $r_s$=0.7876 ($P$<.001; Figure 13B). Daytime (6 AM-10 PM) variability showed a median PPG value of 3.6000 (IQR 0.8667-11.4667) versus 0.0000 (IQR 0.0000-5.600) for 24-hour Holter monitoring, with MAE=4.8425 ($P$<.001) and $r_s$=0.7659 ($P$<.001; Figure 13C), whereas nighttime (10 PM-6 AM) variability showed a median PPG value of 1.000 (IQR 0.0000-6.0000) versus 0.0000 (IQR 0.0000-0.2857) for 24-hour Holter monitoring, with MAE=2.6171 ($P$<.001) and $r_s$=0.6712 ($P$<.001; Figure 13D). Mean real variability measurements yielded a median value of 4.5282 (IQR 0.9031-12.4119) for PPG versus 0.3499 (IQR 0.0000-8.9325) for 24-hour Holter monitoring, with MAE=4.6436 ($P$<.001) and $r_s$=0.8127 ($P$<.001; Figure 13E). Mean AF variation assessment yielded a median PPG value of 0.0466 (IQR –0.2987 to 1.3361) versus –0.0142 (IQR –0.1249 to 0.0654) for 24-hour Holter monitoring, with MAE=0.3893 ($P$=.27) and $r_s$=0.7246 ($P$<.001; Figure 13F).

**Figure 13.** Scatter plot and regression line of AF burden estimated by PPG compared to AF burden estimated by ECG. Each dot represents the AF burden by each method for each patient with AF. Dashed line is the regression line. Solid gray line is line of equality. AF: atrial fibrillation; ECG: electrocardiogram; PPG: photoplethysmography. VOM: variation of the mean (VOM=[ | average AF episode at day−average AF episodes at night | /average 24-h AF episodes] ×100%).




### *Metric M5 Evaluation*

For the proportion of AF episodes with a ventricular rate of >120 bpm to total monitored episodes, PPG measurements showed a median of 0.0007 (IQR 0.0007-0.0027) compared to 0.0000 (IQR 0.0000-0.0070) for 24-hour Holter monitoring, with model performance metrics of MAE=0.0151 ($P$=.76) and $r_s$=0.3435 ($P$<.001). Similarly, for the proportion of time in AF with a ventricular rate of >120 bpm to total monitoring time, the median PPG value was 0.0007 (IQR 0.0007-0.0027) versus 0.0000 (IQR 0.0000-0.0070) for 24-hour Holter monitoring, yielding an identical model performance (MAE=0.0151 and $P$=.76; $r_s$=0.3435 and $P$<.001). Table 6 summarizes the quantification performance of the AF burden model.

**Table .** Photoplethysmography (PPG)–monitored atrial fibrillation (AF) progression features compared to 24-hour electrocardiogram (ECG) monitoring (N=145).

| | PPG, median (IQR) | ECG, median (IQR) | MAE[a] | *P* value | $r_s$ [b] | *P* value |
|---|---|---|---|---|---|---|
| Duration feature | | | | | | |
| Ratio of AF duration to total monitoring time (%) | 0.0402 (0.0086 to 0.8671) | 0.0020 (0.0086 to 0.8671) | 0.0400 | .008 | 0.8788 | <.001 |
| Ratio of AF episodes lasting over 6 min to total monitoring time (%) | 0.0093 (0.0000 to 0.7330) | 0.0000 (0.0000 to 1.0000) | 0.0582 | .44 | 0.9223 | <.001 |
| Ratio of AF episodes lasting over 1 h to total monitoring time (%) | 0.0000 (0.0000 to 0.2947) | 0.0000 (0.0000 to 1.0000) | 0.1204 | .04 | 0.9293 | <.001 |
| Number feature | | | | | | |
| Ratio of AF episodes to total measurements (%) | 0.0033 (0.0009 to 0.0056) | 0.0001 (0.0000 to 0.0008) | 0.0032 | <.001 | –0.0807 | .33 |
| Ratio of AF episodes lasting over 6 min to total measurements (%) | 0.0010 (0.0007 to 0.0028) | 0.0000 (0.0000 to 0.0007) | 0.0015 | <.001 | 0.2360 | .004 |
| Ratio of AF episodes lasting over 1 h to total measurements (%) | 0.0004 (0.0000 to 0.0009) | 0.0000 (0.0000 to 0.0007) | 0.0005 | <.001 | 0.7092 | <.001 |
| Aggregation feature | | | | | | |
| AF density | 0.2700 (0.1700 to 0.4725) | 0.0000 (0.0000 to 0.5900) | 0.1725 | .13 | 0.6576 | <.001 |
| Circadian rhythm feature | | | | | | |
| SD per h of AF episodes | 5.3229 (1.0375 to 15.9058) | 0.4390 (0.0000 to 12.8558) | 5.6009 | <.001 | 0.8135 | <.001 |
| Variability of AF episodes | 3.3478 (0.6957 to 9.2174) | 0.1739 (0.0000 to 5.2174) | 3.9967 | <.001 | 0.7876 | <.001 |
| Daytime AF variability | 3.6000 (0.8667 to 11.4667) | 0.0000 (0.0000 to 5.6000) | 4.8425 | <.001 | 0.7659 | <.001 |
| Nighttime AF variability | 1.0000 (0.0000 to 6.0000) | 0.0000 (0.0000 to 0.2857) | 2.6171 | <.001 | 0.6712 | <.001 |
| Average real variability[c] | 4.5282 (0.9031 to 12.4119) | 0.3499 (0.0000 to 8.9325) | 4.6436 | <.001 | 0.8127 | <.001 |
| VOM[d] of AF episodes (%) | 0.0466 (–0.2987 to 1.3361) | –0.0142 (–0.1249 to 0.0654) | 0.3893 | .27 | 0.7246 | <.001 |
| Heart rate feature | | | | | | |
| Ratio of duration of pulse rate of >120 bpm[e] to monitoring time (min) | 0.0007 (0.0000 to 0.0027) | 0.0000 (0.0000 to 0.0070) | 0.0151 | .76 | 0.3435 | <.001 |
| Ratio of number of pulse rates of >120 bpm to total measurements | 0.0007 (0.0000 to 0.0027) | 0.0000 (0.0000 to 0.0070) | 0.0151 | .76 | 0.3435 | <.001 |

[a]MAE: mean absolute error; mae=∑i=1n|pi−gi|npi:PPGgi:ECG.

[b]Spearman rank correlation coefficient.

[c]Average real variability: average AF episodes, SD, and mean absolute deviation per hour were calculated over 24 hours; average real variability was then equal to the average SD and mean absolute deviation over 24 hours.

[d]VOM: variation of the mean ([ ｜ average AF episodes during the day − average AF episodes at night ｜ /average 24-hour AF episodes] × 100%; daytime AF variability: 6 AM-10 PM; nighttime AF variability: 10 PM-6 AM).

[e]bpm: beats per minute.

## Discussion

### Principal Findings

This study confirmed that the PPG-based AF burden model demonstrated strong agreement in tracking AF burden variations, with metric M1 showing optimal performance: for the ratio of AF duration to the total monitoring time, it achieved an MAE of 0.0400 ($P$=.008) and $r_s$ of 0.8788 ($P$<.001); for the ratio of ≥6-minute AF episodes to the total monitoring time, it achieved an MAE of 0.0582 ($P$=.44) and $r_s$ of 0.9233 ($P$<.001); and for the ratio of ≥1-hour AF episodes to the total monitoring time, it achieved an MAE of 0.1204 ($P$=.04) and $r_s$ of 0.9293 ($P$<.001).

The current clinical classification of AF largely relies on symptomatic presentation and ECG findings, categorizing AF into paroxysmal, persistent, long-standing persistent, and permanent types. The management of AF involves $CHA_2DS_2$-VASc (congestive heart failure; hypertension; age of ≥75 years; diabetes mellitus; previous stroke, transient ischemic attack, or thromboembolism; vascular disease; age of 65-74 years; and sex category)–guided thrombotic risk stratification for anticoagulation decisions. However, current classifications inadequately capture AF progression dynamics [9,11]. For instance, in real-world settings, it remains unclear whether frequent, short-duration AF episodes confer a higher thrombotic risk than infrequent but prolonged episodes. Currently, no universally accepted AF burden definition exists, although it is commonly quantified as the percentage of time in AF during monitoring [12]. Alternative definitions incorporate AF episode frequency, density, and temporal variability [2,5,13], reflecting methodological heterogeneity. While implantable loop recorders and 24-hour Holter monitoring are widely used for AF burden assessment, implantable loop recorders are invasive, costly, and may overestimate AF burden by misclassifying atrial arrhythmias as AF [14]. Conversely, 24-hour Holter monitoring underestimates burden due to short duration and tolerability limitations, restricting its utility in daily practice [15]. This study leveraged PPG integrated smartwatches to enable multidimensional ambulatory AF monitoring, innovatively characterizing AF burden across 5 distinct dimensions: temporal duration, episode frequency, density, variability, and RVR. This approach offers novel insights for establishing a PPG-based definition of AF burden.

In recent years, wearable PPG devices have gained traction for AF screening owing to affordability and usability [8,16]. Previous studies have validated PPG's diagnostic accuracy in AF detection [17-20]. Our team's previous research (N=187,912) confirmed the feasibility and utility of PPG-based wristbands and smartwatches in population level AF screening, facilitating early AF identification and intervention [7]. In a large cohort (N=1,187,381; mean follow-up 255 days), 93.6% of PPG suspected AF cases were verified, demonstrating the reliability and accuracy of this approach [21]. Despite PPG's utility in AF screening, its burden monitoring potential is underexplored. Väliaho et al [8] evaluated PPG-based algorithms for continuous AF burden assessment in 173 participants, finding 30-minute intervals as optimal ($F_1$ score=0.95; sensitivity=94.9%; specificity=98.6%) when comparing 10-, 20-, 30-, and 60-minute reporting intervals. Avram et al [22] evaluated a Samsung smartwatch algorithm (5-minute intervals: sensitivity=87.8%, 95% CI 83.6%-91%; specificity=97.4%, 95% CI 97.1%-97.7%; area under the curve=93.3%), showing strong AF burden correlation ($R^2$=0.986). These findings suggest PPG's potential for AF burden quantification, although detection precision for brief AF episodes requires refinement. In this study, we developed a PPG-based AF burden algorithm integrated into smartwatch devices. When evaluated via 1-minute analysis epochs against standard 24-hour Holter monitoring as the reference, the algorithm demonstrated a sensitivity of 91.5% (95% CI 87.9%-95.1%), specificity of 97.2% (95% CI 95.9%-98.5%), and overall accuracy of 95.7% (95% CI 94%-97.4%). Previous studies have reported PPG signal limitations from motion artifacts and other interfering factors, leading to substantial data exclusion. Reissenberger et al [23] reported that approximately 50.7% of PPG signals were classified as noise, necessitating exclusion. To address these challenges, we implemented algorithm optimizations through a novel multiscale fusion approach for AF burden detection, achieving a daily model output rate of 0.86 (95% CI 0.852-0.880) and markedly improving data usability. In this study, metric M1 demonstrated optimal performance in tracking AF duration, consistent with previous findings, showing an MAE of 0.0400 ($P$=.008) and $r_s$ of 0.8788 ($P$<.001) for the ratio of AF duration to total monitoring time and an MAE of 0.1204 ($P$=.04) and $r_s$=0.9293 ($P$<.001) for the ratio of episodes of >1 hour to total monitoring time. While current studies typically use invasive devices or 24-hour Holter monitoring to define AF burden through episode frequency and density, our study pioneered PPG-based AF burden models integrated into smartwatches. Compared with standard 24-hour Holter monitoring, the AF burden model exhibited higher MAE values, along with a weaker $r_s$, in assessing metrics M2 and M5. This may be attributed to PPG signal interference caused by poor pulse waveform quality, motion artifacts, and noise, which could reduce the models' accuracy in evaluating AF episode frequency. However, metric M3 demonstrated good correlation and a weaker MAE, with an MAE of 0.1725 ($P$=.13) and $r_s$=0.6576 ($P$<.001), although further optimization is still needed in the future. Notably, metric M4 demonstrated excellent

agreement with the gold standard in tracking AF variability, providing novel insights for assessing temporal AF fluctuations.

AF's rising prevalence and mortality exacerbate global health burdens, with its association to adverse cardiovascular events including stroke, heart failure, and death potentially linked to AF burden [1,24]. The Catheter Ablation Versus Standard Conventional Treatment in Patients With Left Ventricular Dysfunction and Atrial Fibrillation trial demonstrated that an AF burden threshold of 50% was associated with significant functional and structural changes in cardiomyocytes [25]. There is accumulating evidence suggesting that AF burden is correlated with adverse cardiovascular events, including stroke and heart failure [2,4,26]. A large-scale study (N=39,710) defined AF burden as the ratio of daily AF duration percentage to total monitoring time [13] and the longest single AF episode duration, revealing dose-dependent associations between increasing baseline AF burden and adverse outcomes at the 1- and 3-year follow-ups. A high AF burden may be a significant risk factor for mortality. However, the causal role of AF burden in stroke events and the optimal threshold (ranging from 1 minute to 24 hours) remain unclear. Moreover, the relationship between AF episode frequency, AF density, AF episode variability, and the proportion of RVR AF and adverse cardiovascular events remains unexplored. However, in this study, metrics M1 and M4 demonstrated strong correlation compared to 24-hour Holter monitoring. This provides a novel methodological foundation for future research into the association between PPG-based AF burden detection and adverse cardiovascular events.

While there is evidence suggesting that anticoagulation therapy may be considered based on atrial high rate episode burden [27], the optimal treatment strategy according to specific AF burden thresholds requires further investigation. Our study's metric M1 for temporal AF burden monitoring showed strong correlation with 24-hour Holter monitoring, potentially offering a novel methodology for future large-scale clinical trials investigating anticoagulant guidance by AF burden quantification. Current AF guidelines recommend active rhythm control to alleviate symptoms in patients with AF, with antiarrhythmic drug therapy and more effective catheter ablation techniques demonstrating the capability to prevent AF recurrence, improve symptoms, and reduce AF burden [28-30]. Drexler et al [31] retrospectively demonstrated that early postprocedural low root mean square of successive differences values following pulmonary vein isolation independently predicted AF recurrence (hazard ratio=0.50; $P<.001$), whereas Zhu et al [32] prospectively identified postablation root mean square of successive differences and percentage of successive normal sinus intervals that differ by more than 50 milliseconds as independent predictors of recurrence in 102 patients with paroxysmal AF. Collectively, recent randomized trials support the beneficial impact of dynamic AF burden monitoring and reduction on clinical outcomes [33]. Our study's metric M4 exhibited strong correlation with 24-hour Holter monitoring in assessing AF variability, providing a potential technical foundation for future rhythm management strategies guided by AF burden quantification and prediction of postablation recurrence risk.

There is emerging evidence suggesting that AF burden may contribute to adverse cardiovascular outcomes, including stroke and heart failure, necessitating refined assessment methods for optimized risk prediction. While real-world evaluation of AF burden remains uncertain, incorporating this parameter into clinical decision making could enable more precise risk stratification and therapeutic selection. Our study developed a novel smartwatch-based PPG algorithm showing strong concordance with 24-hour Holter monitoring in long-term AF burden assessment. This innovative model incorporates 5 dimensions—temporal AF duration, episode density, frequency, variability, and proportion with rapid ventricular response—enabling multidimensional AF evaluation. This approach provides new insights for investigating PPG-derived AF burden's relationship with cardiovascular outcomes, potentially informing early intervention strategies, anticoagulation decisions, and rhythm management protocols.

## Limitations

This study has several limitations that should be acknowledged. First, the relatively modest sample size warrants validation in larger, real-world cohorts to ensure generalizability. In addition, PPG technology remains significantly susceptible to motion artifacts and noise interference, particularly during daily activities and hand movements, which may compromise accurate quantification of AF episodes. This technical limitation presents a notable challenge in assessing RVR AF. However, ongoing advancements in wearable device technology and detection algorithms are expected to mitigate these limitations. Second, while current PPG applications are limited to AF screening and cannot provide definitive diagnosis, our preliminary research demonstrated the diagnostic feasibility of integrated single-lead electrocardiograph technology. Future investigations should prioritize large-scale, multicenter randomized controlled trials to systematically evaluate the clinical utility of combining PPG with iECG technology for comprehensive AF burden monitoring.

## Conclusions

A PPG-based AF burden model incorporating dimensions such as duration, frequency, density, variability, and RVR demonstrated good consistency compared to 24-hour Holter monitoring, with optimal tracking performance observed in the duration and variability dimensions. This approach provides technical support for the at-home multidimensional quantification of AF occurrence and progression, offering new insights for defining PPG-based AF burden. However, further optimization and validation are required for the real-world application of this PPG-based AF burden model.

## Acknowledgments

The authors thank the Huawei Heart Health Research Team for their technical support during the research process. They are grateful to Yu Cao, Hongli Zhang, Yubo You, Teng Xu, Shuai Zhao, and Jie Zhou (Huawei Heart Health Research Team) for the development and optimization of the photoplethysmography-based atrial fibrillation burden model. This study was funded by the National Natural Science Foundation of China (grant 82170309) and the Beijing Natural Science Foundation (grant L232117).

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

Hao Wang and YTG are joint corresponding authors and contributed equally to conceptualization, project administration, funding acquisition, resources, supervision, and writing—review and editing. Hong Wang played a role in the acquisition, analysis, and interpretation of the data. Hong Wang drafted the manuscript. Hong Wang, HZ, ZJ, BL, and ZZ performed data preprocessing. All authors approved the final version of the manuscript.

## Conflicts of Interest

None declared.

## References

1. Van Gelder IC, Rienstra M, Bunting KV, et al. 2024 ESC Guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2024 Sep 29;45(36):3314-3414. [doi: 10.1093/eurheartj/ehae176] [Medline: 39210723]
2. Chew DS, Li Z, Steinberg BA, et al. Arrhythmic burden and the risk of cardiovascular outcomes in patients with paroxysmal atrial fibrillation and cardiac implanted electronic devices. Circ Arrhythm Electrophysiol 2022 Feb;15(2):e010304. [doi: 10.1161/CIRCEP.121.010304] [Medline: 35089799]
3. Wong JA, Conen D, Van Gelder IC, et al. Progression of device-detected subclinical atrial fibrillation and the risk of heart failure. J Am Coll Cardiol 2018 Jun 12;71(23):2603-2611. [doi: 10.1016/j.jacc.2018.03.519] [Medline: 29880119]
4. Go AS, Reynolds K, Yang J, et al. Association of burden of atrial fibrillation with risk of ischemic stroke in adults with paroxysmal atrial fibrillation: the KP-RHYTHM study. JAMA Cardiol 2018 Jul 1;3(7):601-608. [doi: 10.1001/jamacardio.2018.1176] [Medline: 29799942]
5. Charitos EI, Stierle U, Ziegler PD, et al. A comprehensive evaluation of rhythm monitoring strategies for the detection of atrial fibrillation recurrence: insights from 647 continuously monitored patients and implications for monitoring after therapeutic interventions. Circulation 2012 Aug 14;126(7):806-814. [doi: 10.1161/CIRCULATIONAHA.112.098079] [Medline: 22824434]
6. Fan YY, Li YG, Li J, et al. Diagnostic performance of a smart device with photoplethysmography technology for atrial fibrillation detection: pilot study (pre-mAFA II registry). JMIR Mhealth Uhealth 2019 Mar 5;7(3):e11437. [doi: 10.2196/11437] [Medline: 30835243]
7. Guo Y, Wang H, Zhang H, et al. Mobile photoplethysmographic technology to detect atrial fibrillation. J Am Coll Cardiol 2019 Nov 12;74(19):2365-2375. [doi: 10.1016/j.jacc.2019.08.019] [Medline: 31487545]
8. Väliaho ES, Lipponen JA, Kuoppa P, et al. Continuous 24-h photoplethysmogram monitoring enables detection of atrial fibrillation. Front Physiol 2022 Jan 4;12:778775. [doi: 10.3389/fphys.2021.778775] [Medline: 35058796]
9. January CT, Wann LS, Calkins H, et al. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society in collaboration with the Society of Thoracic Surgeons. Circulation 2019 Jul 9;140(2):e125-e151. [doi: 10.1161/CIR.0000000000000665] [Medline: 30686041]
10. Guo Y, Wang H, Zhang H, et al. Photoplethysmography-based machine learning approaches for atrial fibrillation prediction: a report from the Huawei heart study. JACC Asia 2021 Dec 21;1(3):399-408. [doi: 10.1016/j.jacasi.2021.09.004] [Medline: 36341222]
11. Tiver KD, Quah J, Lahiri A, Ganesan AN, McGavigan AD. Atrial fibrillation burden: an update-the need for a CHA2DS2-VASc-AFBurden score. Europace 2021 May 21;23(5):665-673. [doi: 10.1093/europace/euaa287] [Medline: 33351904]
12. Joglar JA, Chung MK, Armbruster AL, et al. 2023 ACC/AHA/ACCP/HRS Guideline for the diagnosis and management of atrial fibrillation: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation 2024 Jan 2;149(1):e1-e156. [doi: 10.1161/CIR.0000000000001193] [Medline: 38033089]

13. Peacock J, Stanelle EJ, Johnson LC, et al. Using atrial fibrillation burden trends and machine learning to predict near-term risk of cardiovascular hospitalization. Circ Arrhythm Electrophysiol 2024 Nov;17(11):e012991. [doi: 10.1161/CIRCEP.124.012991] [Medline: 39445440]
14. Passman RS. Monitoring for AF: identifying the burden of atrial fibrillation and assessing post-ablation. J Innov Card Rhythm Manag 2017 Jan 15;8(1):2575-2582. [doi: 10.19102/icrm.2017.080104] [Medline: 32477752]
15. Singer DE, Atlas SJ, Go AS, et al. Atrial fibrillation burden on a 14-day ECG monitor: findings from the GUARD-AF trial screening arm. JACC Clin Electrophysiol 2025 Jan;11(1):110-119. [doi: 10.1016/j.jacep.2024.08.010] [Medline: 39297839]
16. Charitos EI, Ziegler PD, Stierle U, et al. Atrial fibrillation burden estimates derived from intermittent rhythm monitoring are unreliable estimates of the true atrial fibrillation burden. Pacing Clin Electrophysiol 2014 Sep;37(9):1210-1218. [doi: 10.1111/pace.12389] [Medline: 24665972]
17. Bonini N, Vitolo M, Imberti JF, et al. Mobile health technology in atrial fibrillation. Expert Rev Med Devices 2022 Apr;19(4):327-340. [doi: 10.1080/17434440.2022.2070005] [Medline: 35451347]
18. Chen E, Jiang J, Su R, et al. A new smart wristband equipped with an artificial intelligence algorithm to detect atrial fibrillation. Heart Rhythm 2020 May;17(5 Pt B):847-853. [doi: 10.1016/j.hrthm.2020.01.034] [Medline: 32354449]
19. Tison GH, Sanchez JM, Ballinger B, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. JAMA Cardiol 2018 May 1;3(5):409-416. [doi: 10.1001/jamacardio.2018.0136] [Medline: 29562087]
20. Sibomana O, Hakayuwa CM, Obianke A, Gahire H, Munyantore J, Chilala MM. Diagnostic accuracy of ECG smart chest patches versus PPG smartwatches for atrial fibrillation detection: a systematic review and meta-analysis. BMC Cardiovasc Disord 2025 Feb 25;25(1):132. [doi: 10.1186/s12872-025-04582-2] [Medline: 40000931]
21. Guo Y, Guo J, Shi X, et al. Mobile health technology-supported atrial fibrillation screening and integrated care: a report from the mAFA-II trial Long-term Extension Cohort. Eur J Intern Med 2020 Dec;82:105-111. [doi: 10.1016/j.ejim.2020.09.024] [Medline: 33067121]
22. Avram R, Ramsis M, Cristal AD, et al. Validation of an algorithm for continuous monitoring of atrial fibrillation using a consumer smartwatch. Heart Rhythm 2021 Sep;18(9):1482-1490. [doi: 10.1016/j.hrthm.2021.03.044] [Medline: 33838317]
23. Reissenberger P, Serfözö P, Piper D, et al. Determine atrial fibrillation burden with a photoplethysmographic mobile sensor: the atrial fibrillation burden trial: detection and quantification of episodes of atrial fibrillation using a cloud analytics service connected to a wearable with photoplethysmographic sensor. Eur Heart J Digit Health 2023 Oct 6;4(5):402-410. [doi: 10.1093/ehjdh/ztad039] [Medline: 37794868]
24. Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation 2014 Feb 25;129(8):837-847. [doi: 10.1161/CIRCULATIONAHA.113.005119] [Medline: 24345399]
25. Körtl T, Stehle T, Riedl D, et al. Atrial fibrillation burden specifically determines human ventricular cellular remodeling. JACC Clin Electrophysiol 2022 Nov;8(11):1357-1366. [doi: 10.1016/j.jacep.2022.07.016] [Medline: 36424002]
26. Tanawuttiwat T, Lande J, Smeets P, et al. Atrial fibrillation burden and subsequent heart failure events in patients with cardiac resynchronization therapy devices. J Cardiovasc Electrophysiol 2020 Jun;31(6):1519-1526. [doi: 10.1111/jce.14444] [Medline: 32162753]
27. Healey JS, Lopes RD, Granger CB, et al. Apixaban for stroke prevention in subclinical atrial fibrillation. N Engl J Med 2024 Jan 11;390(2):107-117. [doi: 10.1056/NEJMoa2310234] [Medline: 37952132]
28. Mark DB, Anstrom KJ, Sheng S, et al. Effect of catheter ablation vs medical therapy on quality of life among patients with atrial fibrillation: the CABANA randomized clinical trial. JAMA 2019 Apr 2;321(13):1275-1285. [doi: 10.1001/jama.2019.0692] [Medline: 30874716]
29. Kuniss M, Pavlovic N, Velagic V, et al. Cryoballoon ablation vs. antiarrhythmic drugs: first-line therapy for patients with paroxysmal atrial fibrillation. Europace 2021 Jul 18;23(7):1033-1041. [doi: 10.1093/europace/euab029] [Medline: 33728429]
30. Wazni OM, Dandamudi G, Sood N, et al. Cryoballoon ablation as initial therapy for atrial fibrillation. N Engl J Med 2021 Jan 28;384(4):316-324. [doi: 10.1056/NEJMoa2029554] [Medline: 33197158]
31. Drexler M, Blum T, Heinroth KM, et al. Heart rate variability as a predictor of successful catheter-guided pulmonary vein isolation for atrial fibrillation. Herz 2024 Mar;49(2):147-154. [doi: 10.1007/s00059-023-05201-6] [Medline: 37589750]
32. Zhu Z, Wang W, Cheng Y, Wang X, Sun J. The predictive value of heart rate variability indices tested in early period after radiofrequency catheter ablation for the recurrence of atrial fibrillation. J Cardiovasc Electrophysiol 2020 Jun;31(6):1350-1355. [doi: 10.1111/jce.14448] [Medline: 32173930]
33. Linz D, Andrade JG, Arbelo E, et al. Longer and better lives for patients with atrial fibrillation: the 9th AFNET/EHRA consensus conference. Europace 2024 Mar 30;26(4):euae070. [doi: 10.1093/europace/euae070] [Medline: 38591838]


## Abbreviations

**AF:** atrial fibrillation
**bpm:** beats per minute
**ECG:** electrocardiogram
**MAE:** mean absolute error
**mAFA:** Mobile Atrial Fibrillation Application

**pNN50:** percentage of successive normal sinus intervals that differ by more than 50 milliseconds
**PPG:** photoplethysmography
**RVR:** rapid ventricular rate

*Edited by KC Wong; submitted 26.05.25; peer-reviewed by F Lin, G Romiti, PT Lee; revised version received 23.08.25; accepted 04.09.25; published 10.11.25.*

*Please cite as:*
*Wang H, Liu B, Zhang H, Zhang Z, Jin Z, Wang H, Guo YT*
*Photoplethysmography-Based Machine Learning Approaches for Atrial Fibrillation Burden: Algorithm Development and Validation*
*JMIR Cardio 2025;9:e78075*
*URL: https://cardio.jmir.org/2025/1/e78075*
*doi:10.2196/78075*

# Augmenting Engagement in Decentralized Clinical Trials for Atrial Fibrillation: Development and Implementation of a Programmatic Architecture

Toluwa Daniel Omole[1], BS; Andrew Mrkva[2], MSIS; Danielle Ferry[2], MA; Erin Shepherd[2], BA; Jessica Caratelli[2], BA; Noah Davis[2], MPH; Richmond Akatue[2], BS; Timothy Bickmore[3], PhD; Michael K Paasche-Orlow[4], MD, MA, MPH; Jared W Magnani[5], MD, MSc

[1]University of Pittsburgh School of Medicine, 3609 Forbes Avenue, Second floor, Pittsburgh, PA, United States

[2]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[3]Khoury College of Computer Sciences, Northeastern University, Boston, MA, United States

[4]Department of Medicine, Tufts University School of Medicine, Boston, MA, United States

[5]Center for Research on Health Care, Department of Medicine, University of Pittsburgh, 3609 Forbes Avenue, Pittsburgh, PA, United States

**Corresponding Author:**
Jared W Magnani, MD, MSc
Center for Research on Health Care, Department of Medicine, University of Pittsburgh, 3609 Forbes Avenue, Pittsburgh, PA, United States

## Abstract

**Background:** Atrial fibrillation (AF) is a chronic cardiovascular condition that requires long-term adherence to medications and self-monitoring. Clinical trials for AF have had limited diversity by sex, race and ethnicity, and rural residence, thereby compromising the integrity and generalizability of trial findings. Digital technology coupled with remote strategies has the potential to increase recruitment of individuals from underrepresented demographic and geographic populations, resulting in increased trial diversity, and improvement in the generalizability of interventions for complex diseases such as AF.

**Objective:** This study aimed to summarize the architecture of a research program using remote methods to enhance geographic and demographic diversity in mobile health trials to improve medication adherence.

**Methods:** We developed a programmatic architecture to conduct remote recruitment and assessments of individuals with AF in 2 complementary randomized clinical trials, funded by the National Institutes of Health, to test the effectiveness of a smartphone-based relational agent on adherence to oral anticoagulation. The study team engaged individuals with either rural or metropolitan residences receiving care for AF at health care settings who then provided consent, and underwent baseline assessments and randomization during a remotely conducted telephone visit. Participants were randomized to receive the relational agent intervention or control and subsequently received a study smartphone with installed apps by mail. Participants received a telephone-based training session on device and app usage accompanied by a booklet with pictures and instructions accessible for any level of health or digital literacy. The program included remote methods by mail and telephone to promote retention at 4-, 8-, and 12-month visits and incentivized return of the smartphone following study participation. The program demonstrated excellent participant engagement and retention throughout the duration of the clinical trials.

**Results:** The trials enrolled 513 participants, surpassing recruitment goals for the rural (n=270; target n=264) and metropolitan (n=243; target n=240) studies. A total of 62% (319/513) were women; 31% (75/243) of participants in the metropolitan study were African American, Asian, American Indian or Alaskan native or other races or ethnicities, in contrast to 5% (12/270) in the rural study. Among all participants, 56% (286/513) had less than an associate's degree and 44% (225/513) were characterized as having limited health literacy. Intervention recipients receiving the relational agent used the agent median of 95 - 98 (IQR, 56 - 109) days across both studies. Retention exceeded 89% (457/513) at 12 months with study phones used for median 3.3 (IQR, 1 - 5) participants.

**Conclusions:** We report here the development and implementation of a programmatic architecture for the remote conduct of clinical trials. Our program successfully enhanced trial diversity and composition while providing an innovative mobile health intervention for medication adherence in AF. Our methods provide a model for enhanced recruitment and engagement of diverse participants in cardiovascular trials.

**Trial Registration:** Clinicaltrials.gov NCT04076020; https://clinicaltrials.gov/study/NCT04076020 and Clinicaltrials.gov NCT04075994; https://clinicaltrials.gov/study/NCT04075994

*(JMIR Cardio 2025;9:e66436)* doi:10.2196/66436

**KEYWORDS**

atrial fibrillation; rurality; diversity; mobile health intervention; mobile health; mhealth; chronic cardiovascular condition; cardiovascular; cardio; heart; vascular; medication; self-monitoring; digital health; programmatic architecture; effectiveness; smartphone-based; smartphone; telehealth; telemedicine; digital technology; application; digital literacy; clinical trial; cardiovascular trials

## Introduction

Multiple demographic groups have had limited participation in clinical trials, despite relatively high rates of disease burden [1]. Women, racial and ethnic minorities, and people who reside in rural settings have historically been underrepresented in randomized clinical trials testing or evaluating interventions for cardiovascular diseases [2-6]. Causes of such underrepresentation are multifactorial and related to the individual or patient, investigator, and health care system factors [1,7,8]. Atrial fibrillation (AF) is a chronic cardiovascular condition that merits attention because of its high prevalence and the documented disparities in disease detection, treatment, and outcomes [9]. Global and United States prevalence of AF has increased with concomitant rise in clinical adversity, expenditures, and mortality associated with the condition [10-12]. Individuals with AF experience 4-fold higher rates of inpatient care and 5-fold higher days of hospitalization than those without [13]. The estimated health care costs for AF total US $6 - 20 billion annually, which underscores the importance of trial representation of inclusion of populations that may have increased risks of clinical adversity [14].

Social factors are related to disparities in patient care and experience of AF [15]. Racial and ethnic disparities in AF management are evidenced by Black individuals being less likely to receive oral anticoagulation–a mainstay for thromboembolic stroke prevention in AF–than counterparts of White race [16]. Rural residents may experience structural barriers to care, and in turn lower quality care compared to individuals residing in metropolitan settings [15]. Furthermore, AF is a complicated condition with expectations that patients self-monitor for symptoms, adhere consistently to complex therapies like oral anticoagulants, and have awareness about the disease-related complications [17,18]. Clinical trials likewise have potential to assess the contributions of social and structural factors to patient experience and outcomes in a chronic condition such as AF.

Digital and mobile health interventions have multiple advantages for clinical trials to address the challenges described here. In many circumstances, digital technology can obviate geographic barriers and thereby encourage participation by eliminating travel as a geographic barrier. Such an approach may particularly benefit rural individuals who would otherwise be required to travel as well as metropolitan residents who also experience transportation costs and obstacles [19]. Coupling digital interventions with decentralized trial administration has clear potential to augment the geographic and social diversity of clinical trial participants, which can in turn enhance the generalizability of results and generate new biomedical knowledge [20].

Here we present the design and architecture of a remote mobile health program. We describe a program that uses a smartphone-based intervention to augment the self-management of AF. The intervention incorporates a relational agent [21]—an animated health educator that uses synthetic speech and conversational gestures, such as hand movements, gaze shifts, natural pauses, and emphatic facial expressions to simulate face-to-face counseling—in conjunction with a mobile heart rhythm sensor. We describe here the programmatic architecture to conduct remote trials and the resulting augmentation of geographic, ethnic, and racial diversity of participants. Rather than summarize the results of 2 contrasting trials, our objective in this manuscript is to demonstrate a successful strategy to increase the social and geographic diversity of participants in technology-based trials.

## Methods

### Summary of Recruitment

Our program implemented 2 complementary, parallel-arm, and randomized clinical trials with decentralized administration, summarized here and described in further detail elsewhere [18,22,23]. One trial (ClinicalTrials.gov ID NCT04076020), conducted in individuals with a rural Pennsylvania residence as determined using a definition of rural status as designated by the United States Census Bureau, aimed to recruit geographically remote individuals with AF. The second trial (ClinicalTrials.gov ID NCT04075994) focused on recruitment of individuals residing in metropolitan communities with a focus on economically depressed regions of Pittsburgh, Pennsylvania. Both trials prioritized recruitment from populations that have historically had limited representation in clinical trials for AF. Eligibility for participation in either trial included a diagnosis of AF, as confirmed by the electronic health record, and the prescription of oral anticoagulation for the purpose of thromboembolic stroke prevention in the setting of having AF. The rural and metropolitan trials aimed to recruit 264 and 240 participants, respectively, given differences in design and complementary study objectives.

The architecture of this program consisted of entirely remote recruitment, engagement, assessment, and retention. In effect, this process resulted in the absence of in-person contact between participants and study team members. Recruitment for both studies occurred using multiple approaches. Foremost, eligible individuals received an introductory letter cosigned by their physician provider, such as a physician or nurse practitioner to introduce the research study, accompanied by a brochure, contact information, and a stamped postcard to decline participation. Individuals who did not return the postcard within 2 - 4 weeks received a telephone call as described by the letter. Participants also self-referred, having learned about the study

from their physicians or material placed in clinic settings. Those interested in participating underwent telephone-based screening to verify appropriateness for the trial and review of the inclusion and exclusion criteria summarized in, Textbox 1. If eligible, potential participants scheduled a baseline interview and were then mailed the informed consent and materials for the baseline visit.

**Textbox 1.** Inclusion and exclusion criteria.

**Inclusion criteria**

- Adult, age≥18 years.
- Diagnosis of AF, identified from the electronic health record problem list with confirmation by previous electrocardiogram.
- Prescribed use of warfarin or direct-acting oral anticoagulant for thromboembolic stroke prevention.
- English-speaking at a a level appropriate for informed consent and study participation.
- Residence is defined as rural or in metropolitan Pittsburgh, Pennsylvania.
- No plans to relocate within 12 months of enrollment.

**Exclusion criteria**

- Conditions other than AF that require anticoagulation, such as mechanical prosthetic valve, deep vein thrombosis, or pulmonary embolism.
- Previous electrophysiologic treatment for AF, such as a pulmonary vein isolation.
- Heart failure necessitating hospital admission ≤3 months before study inclusion.
- Acute coronary syndrome (defined as at least 2 of the following: chest pain, ischemic electrocardiographic changes, or troponin≥0.1 ng/mL) ≤3 months before study inclusion.
- Untreated hyperthyroidism or ≤3 months euthyroidism before inclusion.
- Foreseen pacemaker, internal cardioverter defibrillator, or cardiac resynchronization therapy
- Cardiac surgery ≤3 months before inclusion.
- Planned cardiac surgery.
- Presence of noncardiovascular conditions likely to be fatal within 12 months (eg, cancer).
- Inability to comprehend the study protocol, defined by failing 3 times to correctly answer a set of questions during consent.
- A medical disorder, condition, or history that would impair the participant's ability to participate or complete the study.

## Baseline Visit and Randomization

The study team obtained telephone-based informed consent at the start of the baseline visit. The informed consent outlined the study schedule, participant burden and compensation, access to the participant's electronic health record, privacy and security protections in place for participant data in the electronic data management system, and provided contact information for the principal investigator of the study. In addition to providing telephone consent, participants returned a signed copy of the informed consent using a preaddressed, stamped envelope that accompanied study materials.

Consenting participants were randomized to intervention or control arms using electronic software [24,25]. Participants were provided with copies of assessments reproduced in 12- to 14-point font to enhance readability and facilitate their administration by telephone with trained assessors. All study materials were provided at a sixth grade reading level to ensure accessibility to participants. Table 1 lists the assessments conducted by study visit for both trials.

**Table .** Summary of assessments at trial time points (baseline 4, 8, and 12 months) in both randomized clinical trials.

| Measure | Baseline | 4 months | 8 months | 12 months |
|---|---|---|---|---|
| Demographics (age, sex, race, and ethnicity) | ✓ | | | |
| Primary physician for AF[a] treatment | ✓ | | | |
| Transportation (car ownership, driver's license, and distance to physician) | ✓ | | | |
| Mobile device proficiency and ownership [26] | ✓ | | | |
| Social and economic (annual household income, education, and marital status) | ✓ | | | |
| Social network and isolation [27] | ✓ | | | |
| Habits (tobacco and alcohol quantity) | ✓ | | | |
| Medications. Total number and schedules | ✓ | | | |
| AF history (duration and previous treatments) | ✓ | | | |
| Health literacy (Newest Vital Sign) [28,29] | ✓ | | | |
| Clinical conditions, comorbidities, and depression (PHQ-8)[c] [30] | | | | |
| PROMIS[d] Self-efficacy [31] | ✓ | ✓ | ✓ | ✓ |
| Quality of life (AFEQT[e] and PROMIS-29) [32,33] | ✓ | ✓ | ✓ | ✓ |
| Telephone Montreal Cognitive Assessment [34] | ✓ | | | ✓ |
| Medication adherence, self-report [35] | ✓ | ✓ | ✓ | ✓ |
| Proportion of days covered [36] | | ✓ | ✓ | ✓ |
| Health care utilization | | ✓ | ✓ | ✓ |
| New AF therapies and treatments[b] | | ✓ | ✓ | ✓ |
| Qualitative interviews (relational agent, WebMD, and Kardia) | | ✓ | | |

[a]AF therapies: pharmacologic or electrical cardioversion, electrophysiologic procedure such as pulmonary vein isolation, or initiation of antiarrhythmic medication.

[b]AF: atrial fibrillation.

[c]PHQ-8: Patient Health Questionnaire.

[d]PROMIS: Patient-reported Outcomes Measurement Information System.

[e]AFEQT: AF Effect on Quality of life.

## Smartphone Training and Intervention Content

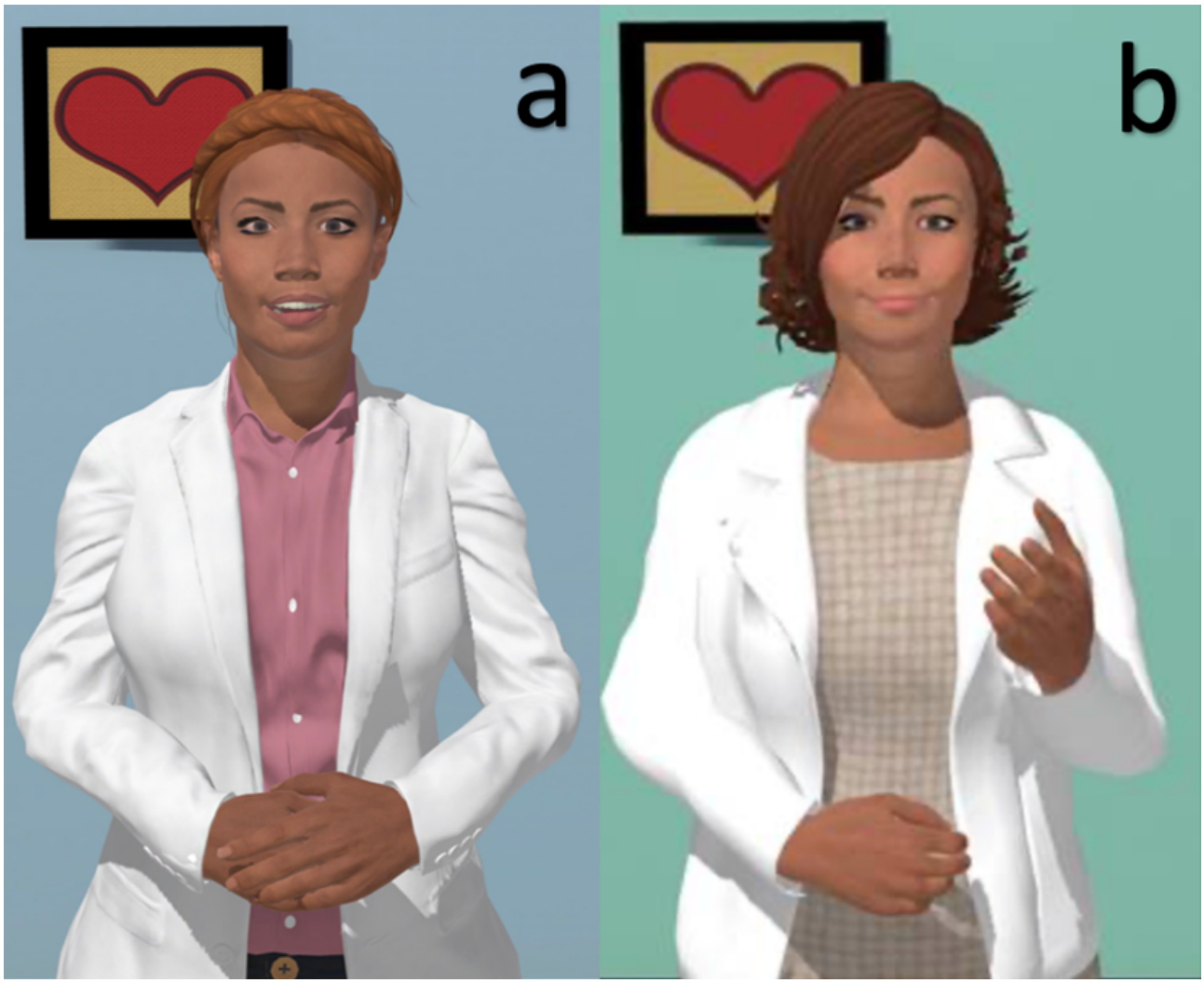
Consented participants received study smartphones accompanied by training on their use and summary guides on smartphone and app operation developed for this study specific to the intervention and control arms. Materials provided to participants in the rural study and randomized to the intervention are provided as an Appendix in Multimedia Appendix 1. Training on smartphone and app use followed a standardized curriculum and concluded when participants reached capacity to operate

the phone at a level appropriate for study participation. Smartphone ownership or digital literacy was consequently not required for trial participation. Smartphones were programmed to enable features such as phone calls and texting without the the capacity for more advanced functions such as downloading apps. Participants randomized to the intervention arms of the trials had the relational agent preinstalled on the smartphone and were provided detailed instructions regarding its use as part of smartphone instruction.

Relational agents are human-computer interventions adapted for multiple settings to facilitate education, problem-solving, and behavioral approaches with patients. Previous work has demonstrated that such agents are accessible to individuals with limited health and digital literacy with varied medical conditions [37-40]. Here, we designed the agent to provide education about AF; problem-solving regarding intentional and unintentional nonadherence to medications; preparation for the medical encounter; and address symptoms common in AF (eg, irregular or rapid heart rates, dyspnea, and chest discomfort). Relational agent content was modified for use in a rural or metropolitan setting. Figure 1 presents the 2 relational agents used in the trials. Intervention participants also received the AliveCor Kardia Mobile device [41] for heart rate and rhythm monitoring, as prompted by the relational agent, to reinforce self-care and the correlation of symptoms with heart rate and rhythm assessments.

Participants randomized to the trials' control arms had the WebMD (WebMD LLC) app preinstalled on the study-provided smartphones. Research assistants encouraged participants to use this application to learn more about AF, its management, and the tracking of medications and symptoms. Participants in the control arms of both trials received an informational session from study team members providing a brief overview of AF and complications such as stroke, heart functioning, and signs of an impending stroke derived from American Heart Association educational materials. To further distinguish the 2 trials, control participants in the metropolitan trial also received the AliveCor Kardia Mobile device with instruction on its use, and guidance to use it as for heart rate and rhythm monitoring. In both studies, heart rate and rhythm obtained using this device was monitored, categorized, and recorded by the study team.

**Figure 1.** Visual representation of the relational agents that were used in the rural trial (Panel A) and metropolitan trial (Panel B) by intervention participants.



### Trial Outcomes

The trials shared the primary outcome to improve adherence to oral anticoagulation for thromboembolic stroke prevention in individuals with AF, as measured by pharmacy claims data using the proportion of days covered [36] (a validated measure to quantify medication adherence using pharmacy claims) and by participant self-report [35]. The secondary outcomes of the trials were (1) health-related quality of life [42], measured using the disease-specific AF Effect on Quality of life [33] and general patient-reported outcomes measurement information system [32] measures; and (2) health care utilization, as measured by days of hospitalization and emergency room visits.

### Participant Timeline and Data Collection

Both intervention and control applications were used for 4 months. Participants were sent a box with a prepaid label for returning the study phone and were informed that their second study payment was tied to a smartphone return. They were allowed to keep the Kardia Mobile device and received instruction from the study team on how to connect the device to their personal smartphone with the caveat that results would no longer be monitored by the study. Participants underwent repeat telephone assessments at 4-, 8-, and 12-months with simultaneous review of the electronic health record for hospitalization events. To assist with interviewer-administered phone-based instrument completion, participants were again mailed the packets of questionnaires summarized in Table 1.

### Remote Engagement and Retention

Given the absence of direct, personal contact, the study developed remote strategies for participant engagement. Participants received regular newsletters also written at 6th-grade reading level for the duration of the study that provided additional education about the studies and updates. In addition, the team mailed birthday cards to participants yearly throughout the duration of the study. Finally, study participants were offered the opportunity to participate in qualitative assessments to further share their experience of AF [43,44]. These sessions were conducted by experienced qualitative researchers using remote video conferencing software.

### Ethical Considerations

The trials described here were registered in clinicaltrials.gov with registration numbers NCT04076020 and NCT04075994 and were approved by the University of Pittsburgh institutional review board. All research participants provided informed consent that allows for secondary analyses without additional consent. This manuscript used solely deidentified data. Participants received compensation up to US $150 for participation across the 4 study visits.

## Results

The rural study enrolled 270 participants while the metropolitan study enrolled 243 participants, in both instances surpassing enrollment goals. Figure 2 shows the geographic representation of participants according to their metropolitan or rural status.

**Figure 2.** Map of the state of Pennsylvania with metropolitan resident participants in purple and rural participants in orange.

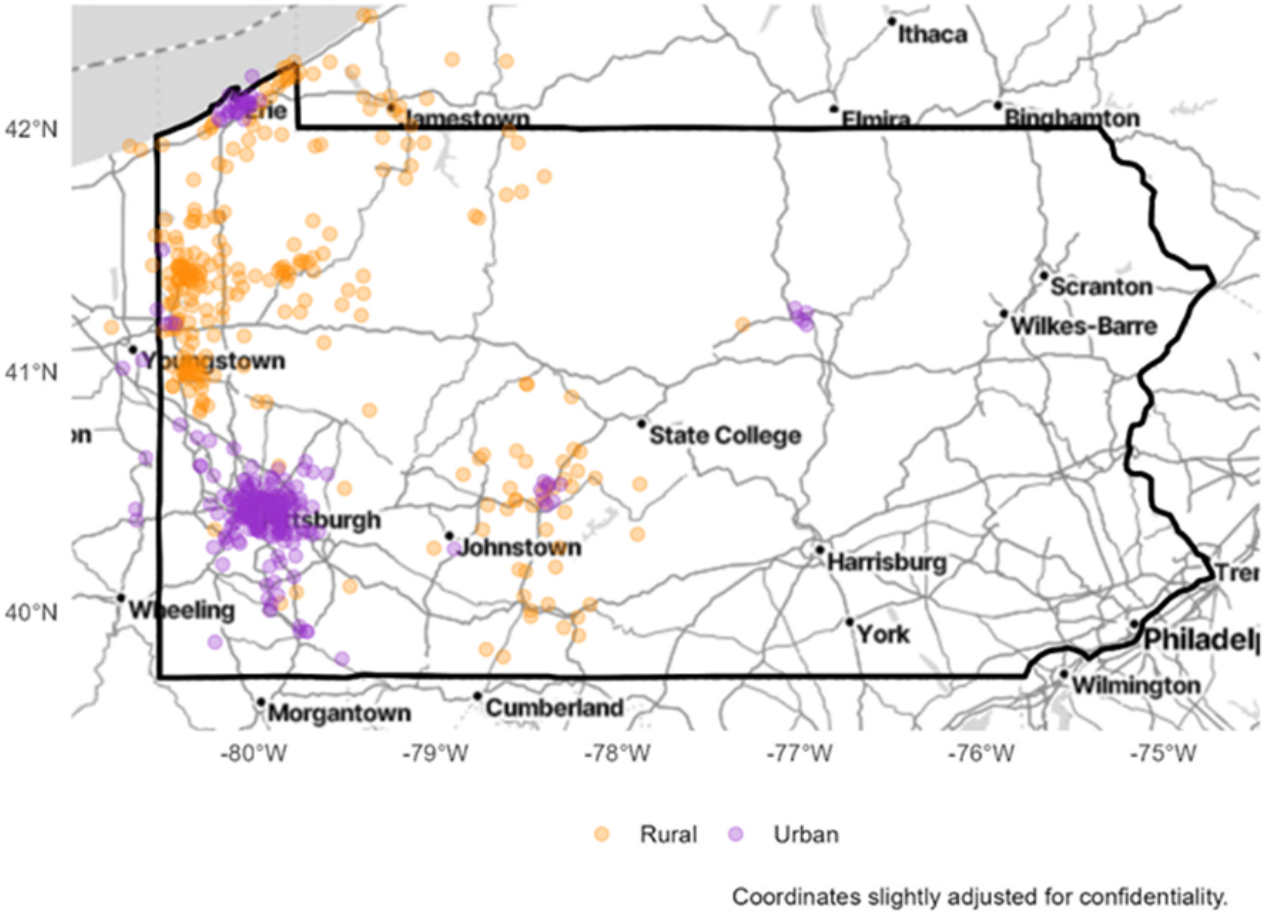



Table 2 summarizes the demographic and social characteristics of participants in both studies. Each study enrolled >60% (319/513) women, consistent with the goal of enrolling individuals with limited participation in clinical trials for AF. In the rural study, 63.7% (172/270) of participants had an educational attainment level less than 4-year college, and 48.5% (131/270) had an annual household income of less than US $50,000/year. The metropolitan study included 30.5% (74/243) individuals of Black race, and 46.9% (114/243) of participants reported an educational attainment of less than 4-year college with 39.5% (96/243) having an annual household income less than US $50,000/year. In both studies, over 41% (225/513) of participants were categorized as having limited health literacy.

**Table .** Baseline characteristics of trial participants in rural Pennsylvania and the metropolitan Pittsburgh, Pennsylvania region.

| Characteristic | Rural Pennsylvania (N=270) | Metropolitan Pittsburgh (N=243) |
|---|---|---|
| Age, years, median (min-max) | 73.1 (40.8-92.2) | 71.6 (29.7-89.6) |
| Years with AF[a], mean (SD) | 6.4 (7.7) | 7.1 (8.1) |
| Sex | | |
| Male | 107 (39.6) | 87 (35.8) |
| Female | 163 (60.4) | 156 (64.2) |
| Race | | |
| White | 257 (95.2) | 161 (66.3) |
| Black | 7 (2.6) | 74 (30.5) |
| Asian | 1 (0.4) | 1 (0.4) |
| American Indian or Alaska Native | 4 (1.5) | — |
| Multiple/Other | — | 5 (2.1) |
| Unknown | 4 (1.6) | 2 (0.8) |
| Ethnicity | | |
| Hispanic/Latino | 5 (1.9) | 0 (0) |
| Not Hispanic/Latino | 261 (96.7) | 240 (98.8) |
| Unknown | 4 (1.4) | 3 (1.2) |
| Education | | |
| High School, vocational, or trade school | 91 (33.7) | 54 (22.2) |
| Vocational or trade School | 37 (13.7) | 18 (7.4) |
| Some college with no degree | 44 (16.3) | 42 (17.3) |
| Associate degree or higher | 98 (36.3.) | 117 (48.1) |
| Unknown | — | 12 (4.9) |
| Employment status | | |
| Employed, full or part-time | 35 (13.0) | 39 (16.0) |
| Retired | 211 (78.2) | 179 (73.7) |
| Other | 24 (8.8) | 25 (10.3) |
| Annual household income (US $) | | |
| <19,999 | 33 (12.2) | 29 (11.9) |
| 20,000 to 34,999 | 52 (19.3) | 35 (14.4) |
| 35,000 to 49,999 | 46 (17.0) | 32 (13.2) |
| 50,000 to 74,999 | 42 (15.6) | 39 (16.0) |
| 75,000 to 99,999 | 27 (10.0) | 27 (11.1 ) |
| ≥100,000 | 30 (11.1) | 34 (14.0) |
| Do not know | 40 (14.8) | 45 (19.4) |
| Type of insurance | | |
| Private | 227 (84.1) | 188 (77.4) |
| Public | 41 (15.2) | 54 (22.2) |
| None | 2 (0.7) | 1 (0.4) |
| Housing | | |
| Ownership | 185 (68.5) | 137 (56.4) |
| Other status | 85 (31.5) | 106 (43.6) |

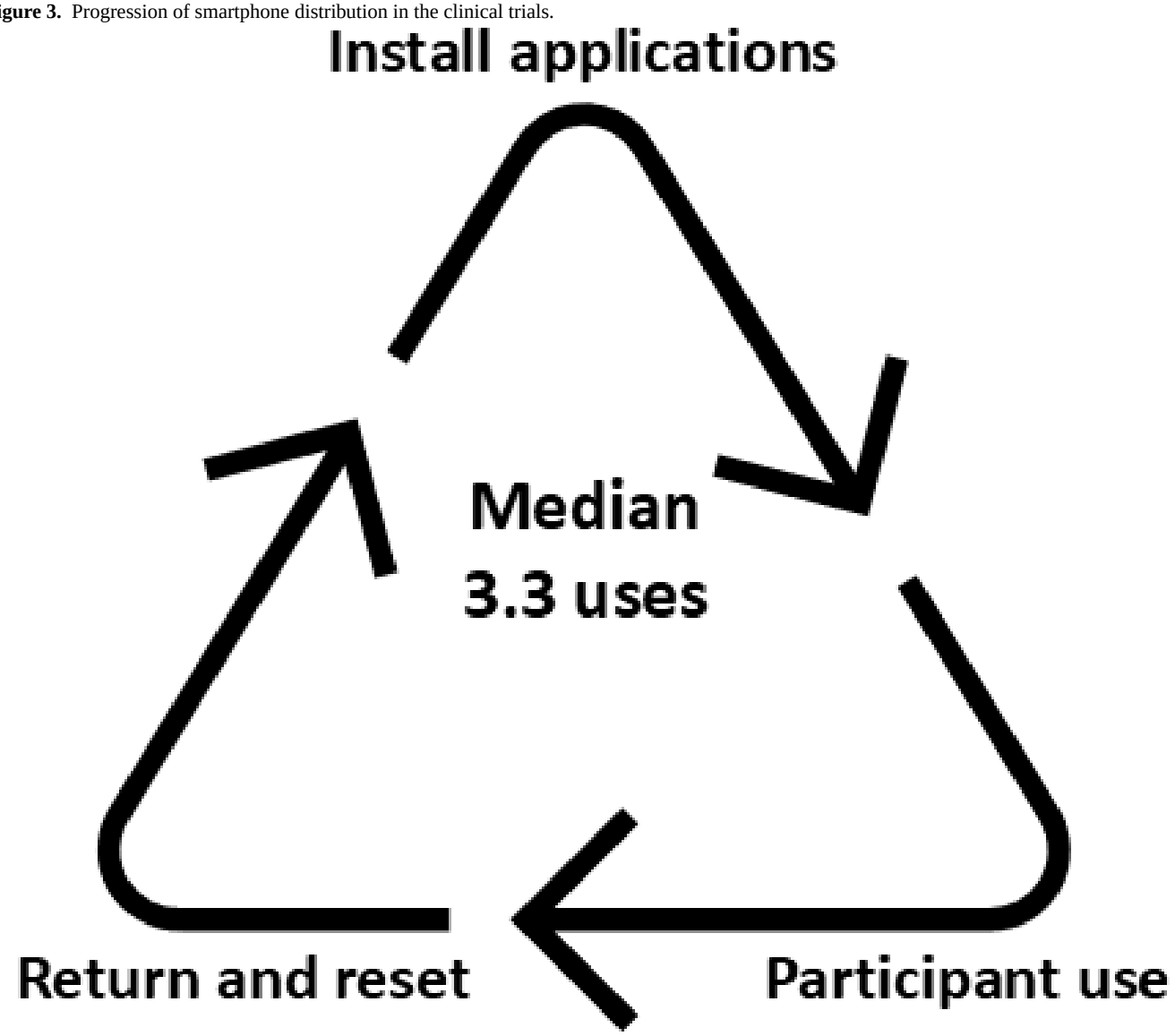
| Characteristic | Rural Pennsylvania (N=270) | Metropolitan Pittsburgh (N=243) |
|---|---|---|
| Marital Status | | |
| Married or living as married | 187 (69.2) | 121 (49.8) |
| Widowed | 49 (18.2) | 45 (18.5) |
| Separated or divorced | 34 (12.6) | 77 (31.7) |
| AF, selection of anticoagulant medication | | |
| Warfarin | 46 (17.0) | 34 (14.0) |
| Direct oral anticoagulant | 225 (83.3) | 209 (86) |
| Health literacy | | |
| Limited health literacy | 125 (46.3) | 100 (41.2) |
| Adequate health literacy | 145 (53.7) | 143 (58.8) |

[a]AF: atrial fibrillation.

Participants randomized to the intervention demonstrated excellent fidelity regarding the use of the relational agent. Rural individuals employed the agent for a median of 101 (IQR, 72 - 110) days of the 120-day trial. Likewise, those randomized to the intervention arm of the metropolitan cohort used the agent for a median of 98 (IQR 58 - 109) days of the 120-day trial. Median days of AliveCor Kardia Mobile device use was 102 (IQR 109 - 123) in the rural trial relative to 95 (IQR 62 - 109) in the metropolitan trial. Out of 270, 239 (88.5%) and 218 (89.7%) participants completed 12-month assessments in the rural and metropolitan trials, respectively.

Phones were returned at 4 months, cleaned, and then reused for additional participants. Figure 3 graphically summarizes the distribution of smartphones. Between the 513 participants of the 2 trials, there were a total of 165 smartphones used with a median use of 3.3 (range 1 - 5) trial participants. In total, 16 phones were lost, stolen, or broken, 5 of which were lost by mail delivery services, not the participant. In addition, 3 phones were lost to participants who withdrew or died without returning the phone to the study.

**Figure 3.** Progression of smartphone distribution in the clinical trials.




## Discussion

### Principal Findings

The decentralized research program described here demonstrates the successful enrollment of rural and metropolitan individuals with AF in clinical trials using a mobile health intervention. Our approach for remote engagement yielded geographic and racial diversity in study participants that exceeds many AF trials. Both trials fulfilled enrollment goals with participants maintaining excellent fidelity to the relational agent intervention and to the AliveCor Kardia Mobile device. Our studies consequently affirm the feasibility of conducting remote, decentralized trials with mobile health interventions, and affirm the demonstrated capacity of decentralized trials to enhance participant diversity [45,46].

### AF and Mobile Health: In the Context of the Literature

AF is a complex syndrome treated with long-term, possibly lifelong, oral anticoagulation for stroke prevention, and managed with subspecialty care and procedures according to professional society guidelines [42]. Our program aimed to address the prominent challenges of health literacy, medication adherence, quality of life, and increased health care utilization that individuals with AF may experience. Our program was informed further by the consistent literature demonstrating the prominent associations of social and structural factors with care processes in AF and its related outcomes [47,48]. Community-based studies, registries, and health services analyses have identified that individuals of non-White race, lower educational attainment, lower income, and residence in neighborhoods with greater social deprivation experience more clinical adversity and limited access to AF-related care than their counterparts [15,16,49-52]. However, social and structural factors are not regularly captured in the conduct of research related to AF, potentially perpetuating disparities by precluding assessment of generalizability to individuals and populations that experience greater social disadvantage. By enrolling individuals with limited education and social resources, our program sought to enhance the access and generalizability of our research program.

Our approach further eliminated geographic and financial barriers to participation. As the intervention was delivered via smartphone, we provided smartphones to participants for the study period along with standardized instruction for their use, thereby eliminating access to contemporary technology as a

barrier to trial participation. Our provision of study materials at a sixth-grade reading level and verbal administration of surveys by staff further diminished health literacy as an obstacle to engagement. Telephone-based visits likewise reduced geographic distance and travel as obstacles to participation. A further iteration of our programmatic design may include video-based visits and additional phenotypic characterization of study participants. Finally, as the literature documents both the increased adversity in women with AF accompanied by diminished participation in clinical trials [53-55], we emphasized the recruitment of women, achieving >60% (319/513) enrollment of women in both trials.

### Promise and Pitfalls of Trials Using Digital Interventions

Remote trials have promise to implement novel digital interventions, such as the relational agent used here. Advantages include the provision of patient-centered education, relevant monitoring, availability, and increased attention and feedback to promote self-care. A meta-analysis determined that digital interventions have the potential to increase medication adherence–the primary outcome of our studies–by 10% (95% CI 1.00 - 1.22) [56].

Concerns for the implementation of digital trials include technological literacy, access to services, and sustainability. Implementation of digital interventions necessitates attention to digital literacy, addressed here by the provision of standardized education and staff support regarding smartphone and device use. Increased dependence on technologies has necessitated the use of digital devices for communication and health maintenance but challenges approximately 20% of individuals reported limited digital literacy in one convenience sample [57]. In addition, the provision of digital technologies requires infrastructure for their effective use. Persistent disparities in Broadband access and coverage present an additional obstacle to the effective implementation of trials using digital technologies. Our study, conducted in metropolitan and rural Pennsylvania, benefited from most participants having adequate cellular coverage and access to the relational agent not being dependent on connectivity. Despite providing access to smartphones, several participants experienced challenges during the trial such as software updates and complications during use as is typical for mobile health trials.

The provision of smartphones to participants eliminated technology access as a barrier to participation. However, we recognize that such an approach would be challenging to sustain beyond the duration of the clinical trials described here. More sustained deployment of mobile health intervention requires assessment of (1) a budget impact, to appreciate the long-term costs (including technology infrastructure and maintenance) and savings of such an intervention and (2) further assessment of the facilitators and barriers that inform the implementation process of the intervention. The next steps of our program include evaluating the implementation process and ascertaining its cost-effectiveness.

### Strengths and Limitations

We recognize several strengths of our programs. We conducted 2 decentralized trials that used a digital health intervention, exceeding recruitment goals in rural and metropolitan settings. Intervention participants demonstrated excellent fidelity with use of the relational agent. Our program also has important, noteworthy limitations that we consider foremost as pertinent to generalizability. First, we recognize rurality as highly heterogeneous and expect that our cohort of rural individuals is not representative of those in other rural contexts. Second, the rural trial was primarily White race, reflecting the region's demographic composition, but again limiting the generalizability of our findings to more racially and ethnically diverse populations. In contrast, the metropolitan study recruited 30% (74/243) individuals of Black race and hence demonstrated greater racial diversity. Third, other settings may benefit from relational agents that are tailored for regional factors such as culture, traditions, digital services, and social and structural factors. We recognize the expansion of agent content as a priority for its implementation in other settings. Together, location, demographic composition, and relational agent design contribute to the limited generalizability of our trials. Finally, the remote design and conduct of assessments by telephone, albeit eliminating multiple obstacles, may be accompanied by a decreased opportunity for more extensive participant characterization and assessments. Remote trial investigators must balance the potential to eliminate participation barriers with the capacity to obtain more robust participant phenotyping and measurements.

### Conclusions

We developed a decentralized, remote research program using a digital intervention. We successfully recruited and enrolled diverse participants that contrast with the relative geographic and social homogeneity of many clinical trials for AF. We intend for our program to provide a roadmap for attaining diverse study participation in digital interventions in decentralized clinical trials for chronic cardiovascular and noncardiovascular diseases.

## Acknowledgments

The authors thank the individuals who participated in this research program. This work was supported by the National Institutes of Health (Awards R01HL143010, R33HL144669, and K24HL160527).

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

## Authors' Contributions

TDO was responsible for investigation and writing–original draft preparation. AM contributed to data curation, formal analysis, investigation, software development, validation, writing–review, and editing. DF provided project administration and supervision and participated in writing–review and editing. ES, JC, and ND contributed to the investigation and writing–review and editing. RA was involved in the investigation and writing– the original draft preparation. TB contributed to conceptualization, funding acquisition, investigation, writing–review, and editing. MK P-O was responsible for conceptualization, funding acquisition, investigation, writing–review, and editing. JWM contributed to conceptualization, data curation, funding acquisition, investigation, methodology, project administration, supervision, writing–review, and editing.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of instructions for smartphone and app use accompanied by standardized instruction.
[PDF File, 6242 KB - cardio_v9i1e66436_app1.pdf ]

## References


1. Michos ED, Reddy TK, Gulati M, et al. Improving the enrollment of women and racially/ethnically diverse populations in cardiovascular clinical trials: An ASPC practice statement. Am J Prev Cardiol 2021 Dec;8(100250):100250. [doi: 10.1016/j.ajpc.2021.100250] [Medline: 34485967]
2. Tahhan AS, Vaduganathan M, Greene SJ, et al. Enrollment of older patients, women, and racial/ethnic minority groups in contemporary acute coronary syndrome clinical trials: a systematic review. JAMA Cardiol 2020 Jun 1;5(6):714-722. [doi: 10.1001/jamacardio.2020.0359] [Medline: 32211813]
3. Khan MS, Shahid I, Siddiqi TJ, et al. Ten-year trends in enrollment of women and minorities in pivotal trials supporting recent US food and drug administration approval of novel cardiometabolic drugs. J Am Heart Assoc 2020 Jun 2;9(11):e015594. [doi: 10.1161/JAHA.119.015594] [Medline: 32427023]
4. Breathett K, Sims M, Gross M, et al. Cardiovascular health in American Indians and Alaska Natives: a scientific statement from the american heart association. Circulation 2020 Jun 23;141(25):e948-e959. [doi: 10.1161/CIR.0000000000000773] [Medline: 32460555]
5. Liu KA, Mager NAD. Women's involvement in clinical trials: historical perspective and future implications. Pharm Pract (Granada) 2016;14(1):708. [doi: 10.18549/PharmPract.2016.01.708] [Medline: 27011778]
6. Khan SU, Khan MZ, Raghu Subramanian C, et al. Participation of women and older participants in randomized clinical trials of lipid-lowering therapies: a systematic review. JAMA Netw Open 2020 May 1;3(5):e205202. [doi: 10.1001/jamanetworkopen.2020.5202] [Medline: 32437574]
7. Oh SS, Galanter J, Thakur N, et al. Diversity in clinical and biomedical research: a promise yet to be fulfilled. PLoS Med 2015 Dec;12(12):e1001918. [doi: 10.1371/journal.pmed.1001918] [Medline: 26671224]
8. Aggarwal NR, Patel HN, Mehta LS, et al. Sex differences in ischemic heart disease: advances, obstacles, and next steps. Circ Cardiovasc Qual Outcomes 2018 Feb;11(2):29449443. [doi: 10.1161/CIRCOUTCOMES.117.004437] [Medline: 29449443]
9. Tertulien T, Magnani JW, Essien UR. Racial and ethnic representation in atrial fibrillation trials: CABANA and beyond. J Am Coll Cardiol 2021 Nov 9;78(19):e163-e164. [doi: 10.1016/j.jacc.2021.08.054] [Medline: 34736573]
10. Zimetbaum P. Atrial fibrillation. Ann Intern Med 2017 Mar 7;166(5):ITC33-ITC48. [doi: 10.7326/AITC201703070] [Medline: 28265666]
11. Kornej J, Börschel CS, Benjamin EJ, Schnabel RB. Epidemiology of atrial fibrillation in the 21st century: novel methods and new insights. Circ Res 2020 Jun 19;127(1):4-20. [doi: 10.1161/CIRCRESAHA.120.316340] [Medline: 32716709]
12. Jiang S, Seslar SP, Sloan LA, Hansen RN. Health care resource utilization and costs associated with atrial fibrillation and rural-urban disparities. J Manag Care Spec Pharm 2022 Nov;28(11):1321-1330. [doi: 10.18553/jmcp.2022.28.11.1321] [Medline: 36282926]
13. Deshmukh A, Iglesias M, Khanna R, Beaulieu T. Healthcare utilization and costs associated with a diagnosis of incident atrial fibrillation. Heart Rhythm O2 2022 Oct;3(5):577-586. [doi: 10.1016/j.hroo.2022.07.010] [Medline: 36340482]
14. Kim MH, Johnston SS, Chu BC, Dalal MR, Schulman KL. Estimation of total incremental health care costs in patients with atrial fibrillation in the United States. Circ Cardiovasc Qual Outcomes 2011 May;4(3):313-320. [doi: 10.1161/CIRCOUTCOMES.110.958165] [Medline: 21540439]
15. Essien UR, Kornej J, Johnson AE, Schulson LB, Benjamin EJ, Magnani JW. Social determinants of atrial fibrillation. Nat Rev Cardiol 2021 Nov;18(11):763-773. [doi: 10.1038/s41569-021-00561-0] [Medline: 34079095]
16. Essien UR, Kim N, Hausmann LRM, et al. Disparities in anticoagulant therapy initiation for incident atrial fibrillation by Race/Ethnicity among patients in the Veterans Health Administration System. JAMA Netw Open 2021 Jul 1;4(7):e2114234. [doi: 10.1001/jamanetworkopen.2021.14234] [Medline: 34319358]

17. Riegel B, Moser DK, Buck HG, et al. Self-care for the prevention and management of cardiovascular disease and stroke: a scientific statement for healthcare professionals from the American Heart Association. J Am Heart Assoc 2017 Aug 31;6(9):28860232. [doi: 10.1161/JAHA.117.006997] [Medline: 28860232]
18. Althouse AD, Abebe KZ, Paasche-Orlow MK, et al. Design, rationale, and baseline characteristics of a randomized controlled trial evaluating a mobile relational agent to enhance atrial fibrillation self-care. Contemp Clin Trials 2023 Jan;124:107015. [doi: 10.1016/j.cct.2022.107015] [Medline: 36402276]
19. Dorsey ER, Kluger B, Lipset CH. The new normal in clinical trials: decentralized studies. Ann Neurol 2020 Nov;88(5):863-866. [doi: 10.1002/ana.25892] [Medline: 32869367]
20. Schwartz AL, Alsan M, Morris AA, Halpern SD. Why diverse clinical trial participation matters. N Engl J Med 2023 Apr 6;388(14):1252-1254. [doi: 10.1056/NEJMp2215609] [Medline: 37017480]
21. Bickmore T, Schulman D, Yin L. Maintaining engagement in long-term interventions with relational agents. Appl Artif Intell 2010 Jul 1;24(6):648-666. [doi: 10.1080/08839514.2010.492259] [Medline: 21318052]
22. Magnani JW, Ferry D, Swabe G, et al. Design and rationale of the mobile health intervention for rural atrial fibrillation. Am Heart J 2022 Oct;252(16-25):16-25. [doi: 10.1016/j.ahj.2022.05.023] [Medline: 35691371]
23. Magnani JW, Ferry D, Swabe G, et al. Rurality and atrial fibrillation: a pathway to virtual engagement and clinical trial recruitment in response to COVID-19. Am Heart J Plus 2021 Mar;3:34151310. [doi: 10.1016/j.ahjo.2021.100017] [Medline: 34151310]
24. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95(103208):103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
25. blockrand SG. Randomization for block random clinical trials. URL: https://CRAN.R-project.org/package=blockrand [accessed 2025-05-07]
26. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults: the mobile device proficiency questionnaire. J Appl Gerontol 2018 Feb;37(2):131-156. [doi: 10.1177/0733464816642582] [Medline: 27255686]
27. Berkman LF, Syme SL. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. Am J Epidemiol 1979 Feb;109(2):186-204. [doi: 10.1093/oxfordjournals.aje.a112674] [Medline: 425958]
28. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594. [Medline: 15343421]
29. Weiss BD, Mays MZ, Martz W, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522. [doi: 10.1370/afm.405] [Medline: 16338915]
30. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
31. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-262. [Medline: 11769298]
32. Cella D, Riley W, Stone A, et al. The patient reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194. [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
33. Spertus J, Dorian P, Bubien R, et al. Development and validation of the atrial fibrillation effect on quality-of-life (AFEQT) Questionnaire in patients with atrial fibrillation. Circ Arrhythm Electrophysiol 2011 Feb;4(1):15-25. [doi: 10.1161/CIRCEP.110.958033] [Medline: 21160035]
34. Katz MJ, Wang C, Nester CO, et al. T-MoCA: A valid phone screen for cognitive impairment in diverse community samples. Alzheimers Dement (Amst) 2021;13(1):e12144. [doi: 10.1002/dad2.12144] [Medline: 33598528]
35. Voils CI, Maciejewski ML, Hoyle RH, et al. Initial validation of a self-report measure of the extent of and reasons for medication nonadherence. Med Care 2012 Dec;50(12):1013-1019. [doi: 10.1097/MLR.0b013e318269e121] [Medline: 22922431]
36. Peterson AM, Nau DP, Cramer JA, Benner J, Gwadry-Sridhar F, Nichol M. A checklist for medication compliance and persistence studies using retrospective databases. Value Health 2007;10(1):3-12. [doi: 10.1111/j.1524-4733.2006.00139.x] [Medline: 17261111]
37. Bickmore TW, Silliman RA, Nelson K, et al. A randomized controlled trial of an automated exercise coach for older adults. J Am Geriatr Soc 2013 Oct;61(10):1676-1683. [doi: 10.1111/jgs.12449] [Medline: 24001030]
38. Bickmore TW, Utami D, Matsuyama R, Paasche-Orlow MK. Improving access to online health information with conversational agents: a randomized controlled experiment. J Med Internet Res 2016 Jan 4;18(1):e1. [doi: 10.2196/jmir.5239] [Medline: 26728964]
39. King AC, Bickmore TW, Campero MI, Pruitt LA, Yin JL. Employing virtual advisors in preventive care for underserved communities: results from the COMPASS study. J Health Commun 2013;18(12):1449-1464. [doi: 10.1080/10810730.2013.798374] [Medline: 23941610]
40. King AC, Campero MI, Sheats JL, et al. Effects of counseling by peer human advisors vs computers to increase walking in underserved populations: The COMPASS randomized clinical trial. JAMA Intern Med 2020 Nov 1;180(11):1481-1490. [doi: 10.1001/jamainternmed.2020.4143] [Medline: 32986075]

41. Wegner FK, Kochhäuser S, Ellermann C, et al. Prospective blinded evaluation of the smartphone-based AliveCor Kardia ECG monitor for atrial fibrillation detection: The PEAK-AF study. Eur J Intern Med 2020 Mar;73(72-5):72-75. [doi: 10.1016/j.ejim.2019.11.018] [Medline: 31806411]
42. Writing Committee Members, Joglar JA, Chung MK, Armbruster AL, Benjamin EJ, Chyou JY, et al. 2023 ACC/AHA/ACCP/HRS Guideline for the diagnosis and management of atrial fibrillation: a report of the American college of cardiology/American Heart Association Joint Committee on clinical practice guidelines. J Am Coll Cardiol 2024 Jan;83(1):109-279. [doi: 10.1016/j.jacc.2023.08.017]
43. Mann H, Johnson AE, Ferry D, et al. A qualitative crossroads of rhythm and race: Black patients' experiences living with atrial fibrillation. Am Heart J Plus 2023 Apr;28:37181157. [doi: 10.1016/j.ahjo.2023.100293] [Medline: 37181157]
44. Mann HK, Streiff M, Schultz KC, et al. Rurality and atrial fibrillation: patient perceptions of barriers and facilitators to care. J Am Heart Assoc 2023 Nov 7;12(21):e031152. [doi: 10.1161/JAHA.123.031152] [Medline: 37889198]
45. Mayfield JJ, Chatterjee NA, Noseworthy PA, et al. Implementation of a fully remote randomized clinical trial with cardiac monitoring. Commun Med (Lond) 2021;1(62):62. [doi: 10.1038/s43856-021-00052-w] [Medline: 35604806]
46. Stewart J, Krows ML, Schaafsma TT, et al. Comparison of racial, ethnic, and geographic location diversity of participants enrolled in clinic-based vs 2 remote COVID-19 clinical trials. JAMA Netw Open 2022 Feb 1;5(2):e2148325. [doi: 10.1001/jamanetworkopen.2021.48325] [Medline: 35157053]
47. Martin SS, Aday AW, Almarzooq ZI, et al. 2024 heart disease and stroke statistics: a report of US and global data from the American Heart Association. Circulation 2024 Feb 20;149(8):e347-e913. [doi: 10.1161/CIR.0000000000001209] [Medline: 38264914]
48. Benjamin EJ, Thomas KL, Go AS, et al. Transforming atrial fibrillation research to integrate social determinants of health: a National Heart, Lung, and Blood Institute workshop report. JAMA Cardiol 2023 Feb 1;8(2):182-191. [doi: 10.1001/jamacardio.2022.4091] [Medline: 36478155]
49. Essien UR, Chiswell K, Kaltenbach LA, et al. Association of race and ethnicity with oral anticoagulation and associated outcomes in patients with atrial fibrillation: findings from the get with the guidelines-atrial fibrillation registry. JAMA Cardiol 2022 Dec 1;7(12):1207-1217. [doi: 10.1001/jamacardio.2022.3704] [Medline: 36287545]
50. Magnani JW, Norby FL, Agarwal SK, et al. Racial differences in atrial fibrillation-related cardiovascular disease and mortality: the atherosclerosis risk in communities (ARIC) study. JAMA Cardiol 2016 Jul 1;1(4):433-441. [doi: 10.1001/jamacardio.2016.1025] [Medline: 27438320]
51. Tertulien T, Chen Y, Althouse AD, Essien UR, Johnson AE, Magnani JW. Association of income and educational attainment in hospitalization events in atrial fibrillation. Am J Prev Cardiol 2021 Sep;7:100201. [doi: 10.1016/j.ajpc.2021.100201] [Medline: 34611640]
52. Abdel-Qadir H, Akioyamen LE, Fang J, et al. Association of neighborhood-level material deprivation with atrial fibrillation care in a single-payer health care system: a population-based cohort study. Circulation 2022 Jul 19;146(3):159-171. [doi: 10.1161/CIRCULATIONAHA.122.058949] [Medline: 35678171]
53. Emdin CA, Wong CX, Hsiao AJ, et al. Atrial fibrillation as risk factor for cardiovascular disease and death in women compared with men: systematic review and meta-analysis of cohort studies. BMJ 2016 Jan 19;532:h7013. [doi: 10.1136/bmj.h7013] [Medline: 26786546]
54. Piccini JP, Simon DN, Steinberg BA, et al. Differences in clinical and functional outcomes of atrial fibrillation in women and men: two-year results from the ORBIT-AF registry. JAMA Cardiol 2016 Jun 1;1(3):282-291. [doi: 10.1001/jamacardio.2016.0529] [Medline: 27438106]
55. Khan SU, Raghu Subramanian C, Khan MZ, et al. Association of women authors with women enrollment in clinical trials of atrial fibrillation. J Am Heart Assoc 2022 Mar;11(5):e024233. [doi: 10.1161/JAHA.121.024233] [Medline: 35191324]
56. Akinosun AS, Polson R, Diaz-Skeete Y, et al. Digital technology interventions for risk factor modification in patients with cardiovascular disease: systematic review and meta-analysis. JMIR Mhealth Uhealth 2021 Mar 3;9(3):e21061. [doi: 10.2196/21061] [Medline: 33656444]
57. Jongebloed H, Anderson K, Winter N, et al. The digital divide in rural and regional communities: a survey on the use of digital health technology and implications for supporting technology use. BMC Res Notes 2024 Mar 28;17(1):90. [doi: 10.1186/s13104-024-06687-x] [Medline: 38549176]


## Abbreviations

**AF:** atrial fibrillation

*Edited by A Coristine; submitted 12.09.24; peer-reviewed by B Fadojutimi, S Mitra; revised version received 14.03.25; accepted 06.04.25; published 12.05.25.*

*Please cite as:*
*Omole TD, Mrkva A, Ferry D, Shepherd E, Caratelli J, Davis N, Akatue R, Bickmore T, Paasche-Orlow MK, Magnani JW*
*Augmenting Engagement in Decentralized Clinical Trials for Atrial Fibrillation: Development and Implementation of a Programmatic Architecture*
*JMIR Cardio 2025;9:e66436*
*URL: https://cardio.jmir.org/2025/1/e66436*
*doi:10.2196/66436*

# Serial 12-Lead Electrocardiogram–Based Deep-Learning Model for Hospital Admission Prediction in Emergency Department Cardiac Presentations: Retrospective Cohort Study

Arda Altintepe[1]; Kutsev Bengisu Ozyoruk[2], PhD

[1]Horace Mann School, New York, NY, United States

[2]Department of Radiology and Imaging Sciences, Emory University School of Medicine, Health Sciences Research Building II (HSRB II), 1760 Haygood Dr NE Suite N647, Atlanta, GA, United States

**Corresponding Author:**
Kutsev Bengisu Ozyoruk, PhD
Department of Radiology and Imaging Sciences, Emory University School of Medicine, Health Sciences Research Building II (HSRB II), 1760 Haygood Dr NE Suite N647, Atlanta, GA, United States

## Abstract

**Background:** Emergency department (ED) crowding is often attributed to a slow hospitalization process, leading to reduced quality of care. Predicting early disposition in patients presenting with cardiac issues is challenging: most are ultimately discharged, yet those with a cardiac etiology frequently require hospital admission. Existing scores rely on single-time-point data and often underperform when patient risk evolves during the visit.

**Objective:** This study aimed to develop and validate a real-time deep-learning model that fuses serial 12-lead electrocardiogram (ECG) waveforms with sequential vitals and routinely available clinical data to predict hospital admission early in ED encounters.

**Methods:** We conducted a retrospective cohort study using the Medical Information Mart for Intensive Care (MIMIC) IV, MIMIC-IV Emergency Department module, and MIMIC-IV electrocardiogram module databases. Adults presenting with chest pain, dyspnea, syncope, or presyncope and at least 1 ECG within their ED stay were included. Two evaluation cohorts were defined: all stays with ≥1 ECG (n=30,421) and a subset with ≥2 ECGs during the encounter (n=11,273). To predict hospital admission, we first established 2 baseline models: a tabular model (random forest [RF]) trained on structured clinical variables, including demographics, triage acuity, past medical history, medications, and laboratory results, and an ECG-only model that learned directly from raw 12-lead waveforms. We then developed a multimodal deep-learning model that combined ECGs with sequential vital signs as well as the same static tabular features. All models were restricted to data available during the stay up to the time of the last ECG. Performance was assessed with stratified 5-fold cross-validation using identical splits across models.

**Results:** The multimodal model achieved an area under receiver operating characteristic (AUROC) of 0.911 when trained on all eligible stays. The model predicted disposition after the final ECG was taken, which was a median of 0.3 (IQR 0.2 - 5.3) hours after triage and 4.6 (IQR 2.7 - 7.3) hours before ED departure. Baseline models performed worse: the ECG-only model had an AUROC of 0.852, and the tabular RF had an AUROC of 0.886. In the subset requiring at least 2 ECGs within the stay, ECG-only reached an AUROC of 0.859, and RF, with the longer interval to chart tabular data, reached a higher AUROC of 0.911. The multimodal model had an AUROC of 0.924 and outperformed baselines in each cohort (paired DeLong $P$<.001).

**Conclusions:** Serial ECGs, when integrated with evolving vitals and routine clinical features, enable accurate, early prediction of ED disposition in patients presenting with cardiac issues. This open-source, reproducible framework highlights the potential of multimodal deep learning to streamline ED flow, prioritize higher risk cases, and detect evolving, time-critical pathology.

***(JMIR Cardio 2025;9:e80569)*** doi:10.2196/80569

**KEYWORDS**

emergency department; electrocardiography; deep learning; multimodal machine learning; transfer learning; hospital admission prediction; cardiac presentations; serial ECG; MIMIC-IV; electrocardiogram

## Introduction

Chest pain, dyspnea, and syncope are among the most common emergency department (ED) chief complaints that eventually result in a cardiac diagnosis [1]. They make up around 16 million encounters yearly in the United States, with chest pain accounting for almost 11 million visits a year [2-4]. Despite their cardiac connotation, the majority prove noncardiac: observational series show that 58.7% of chest pain cases are discharged with a noncardiac diagnosis, adjudication of an international dyspnea cohort found cardiac etiology in 47% and noncardiac causes in the remaining 53% of patients, and only

7% - 10% of syncope presentations to ED are ultimately attributed to a cardiac mechanism [5,6]. However, when a cardiac condition is confirmed, hospitalization becomes far more likely: more than 80% of acute heart failure presentations and up to 86% of high-risk syncope cases are admitted, while admission is far less common for patients whose symptoms are ultimately noncardiac [7,8].

ED providers typically rely on an initial assessment that includes history and physical examination, vital signs, cardiac biomarker tests, and clinical risk scores, such as the emergency severity index (ESI) triage level or the history, electrocardiogram (ECG), age, risk factors, and Troponin (HEART; History, ECG, Age, Risk factors, and Troponin) score [2,9]. However, these traditional risk stratification tools have important limitations. Scores such as ESI and HEART are calculated at a single time point and may not fully reflect evolving patient risk. In practice, initial risk stratification for possible patients presenting with cardiac issues can be insufficient, contributing to ED crowding, which is associated with delayed care, mortality, and generally poorer patient outcomes [10]. This motivates the exploration of advanced machine learning (ML) methods that can integrate multiple data sources and time points to improve predictive performance.

Recent studies have shown that ML and deep-learning models can outperform traditional triage and risk scores in predicting outcomes for patients presenting to the ED. These models often use triage data in combination with vitals, lab results, free-text notes, and past medical history (PMH) to successfully predict general hospitalization or specific critical care outcomes such as acute coronary syndrome [9,11-24]. Fewer recent studies include features within the ED stay, including medications administered, lab tests, and early diagnoses [22-24].

One promising avenue is leveraging deep learning to fuse heterogeneous data sources, including sequential time-series data, such as waveforms, for outcome prediction. Many studies have incorporated ECGs into a disposition model; however, they use implied ECG findings indicated within physician notes or a simple flag indicating whether an ECG was abnormal or conducted [9,15-24]. The implementation of waveforms or more advanced ECG features remains unexplored. Patients with chest pain, dyspnea, presyncope, and syncope often undergo ECGs and vital sign measurements over the course of their ED evaluation. Important prognostic information may lie in the trends and changes in these data. Prior work suggests that sequential data modeling can improve the prediction of patient outcomes. For instance, Bouzid et al [25] demonstrated that analyzing serial ECGs can enhance the detection of acute coronary syndromes: in patients with suspected non-ST-elevation myocardial infarction, combining the prehospital ECG with the initial ED ECG and applying an ML classifier improved diagnostic accuracy and an AI-augmented model further boosted performance to an area under receiver operating characteristic (AUROC) score of 0.83.

Our study builds on these advancements by introducing a multimodal deep-learning approach for early prediction of ED disposition in adult patients presenting with cardiac-related complaints. We developed a multimodal deep-learning model that fuses serial ECG waveforms, sequential vital signs, and key clinical features to predict, in real time, whether a chest-pain patient will require hospital admission. The model will predict disposition after the final ECG available has been taken, requiring anywhere from 1 to 6 ECGs. In contrast to prior works that often focus on diagnostic endpoints or use data available only at presentation, we target the practical outcome of patient disposition and leverage data collected during the ED stay, focusing on ECG waveforms, a novel data stream when predicting ED disposition.

The vast majority of previous studies used private hospital data that were not available for public use [9,12-15,17-24]. The Medical Information Mart for Intensive Care IV (MIMIC IV) is a large deidentified dataset of patients admitted to the ED or an intensive care unit at the Beth Israel Deaconess Medical Center in Boston, MA [26-28]. All data used in this project can be found in the MIMIC-IV, MIMIC-IV Emergency Department module (MIMIC-IV-ED), and MIMIC-IV electrocardiogram module (MIMIC-IV-ECG) modules on PhysioNet [26-30].

The aim of this study was to develop and validate a multimodal deep-learning model that integrates ECG waveforms, sequential vital signs, and tabular clinical data to predict hospital admission in real time. We hypothesized that this fusion of modalities would yield more accurate predictions than conventional methods. If successful, our approach could improve early identification of patients with chest pain, dyspnea, and syncope who require admission (or conversely, those who are safe for early discharge), ultimately enhancing ED decision-making, resource use, and patient outcomes.

## Methods

### Study Cohort

The MIMIC-IV-ED module was filtered to obtain a cohort of 82,907 unique patients with at least 1 mapped ECG from the MIMIC-IV-ECG module within the duration of an ED stay [29,30]. ED stays without ECGs were excluded, and 1 stay per patient was retained.

The cohort was then filtered only to include patients whose chief complaints contained keywords relating to presyncope, syncope, dyspnea, and chest pain. After filtering patients with a disposition other than discharge or admission, 30,421 unique patients presenting to the ED with possible cardiac-related symptoms and at least 1 ECG were left in the study (Figure 1). The number of ECGs per stay and the length of stay before the final ECG from which the model predicts disposition were additionally noted as static features.

**Figure 1.** Flowchart of study population. ECG: electrocardiogram; ED: emergency department; MIMIC-IV-ECG: Medical Information Mart for Intensive Care IV electrocardiogram module.

## Feature Extraction

### *ED Lab Results and Medications (Static)*

From MIMIC-IV hospital data, we extracted ED labs and medications recorded after ED arrival and before the final ECG. The lab results included were Troponin T, creatinine, lactate, C-reactive protein, B-type natriuretic peptide, hemoglobin, potassium, magnesium, and white blood cell count. For each lab, we derived 4 features: first value, peak value, an abnormal flag, and a "missing" flag. ED medications were identified through Pyxis Generic Sequence Number codes, which were mapped to Enhanced Therapeutic Classification codes and grouped by their first 6 digits. Each group contributed a binary feature (38 features). All lab tests and medications were incorporated as static features and were not modeled sequentially or imputed, since both missingness and sequence length varied substantially across patients, making static summarization the most consistent approach.

### *PMH (Static)*

We included the disposition of the prior visit and the number of previous stays visible in the MIMIC-IV database. Prior *International Classification of Diseases* diagnoses from a MIMIC hospital visit were binned to Clinical Classifications Software for *International Classification of Diseases, 9th Revision* and Clinical Classifications Software Redefined for *International Statistical Classification of Diseases and Related Health Problems 10th Revision* categories (253 variables). Outpatient prescriptions in the past year were binned to 7 broad Enhanced Therapeutic Classification groups.

### *Other Static Features*

Age, sex, and acuity were included along with arrival mode and chief complaint Boolean operators. We added total counts for ED medications, prior medications, and prior diagnoses.

### *Vitals (Sequential)*

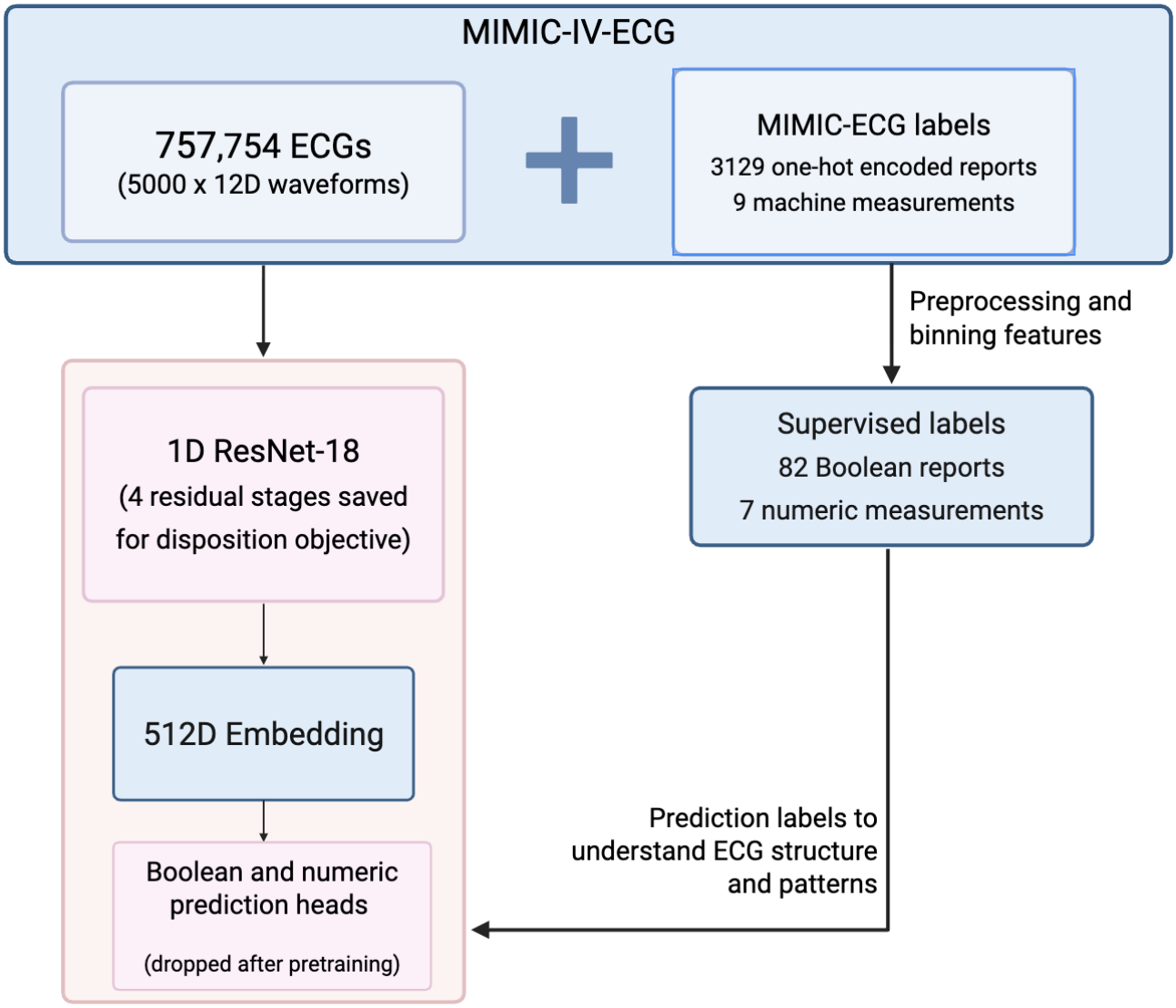
We incorporated temperature, heart rate, respiratory rate, oxygen saturation, systolic and diastolic blood pressure, and pain recorded before the final ECG. All vitals taken at triage were included, as well as vitals drawn throughout the ED stay, with chart time before the last ECG for that stay. All vitals except for pain were nonmissing at triage but contained a large number of missing values when later charted during the stay. Missing vitals, excluding pain, were imputed using a Bayesian Regression formula fit on the training set per fold. Missing flags were created for each vital to note whether the vital was initially missing before imputation. Two additional columns were added, stating whether the patient was sleeping or unable to answer based on the string value of pain.

### *ECG Feature Extraction*

In order to best implement ECGs for the downstream task of predicting hospitalization, an AI model was used to encode a lower dimensional ECG feature vector. Specifically, a one-dimensional residual neural network model (ResNet-18) was chosen due to its ability to learn discriminative temporal and spatial features from high-dimensional ECG signals while avoiding vanishing gradients through skip connections. The network takes in an ECG with 5000 samples (500 Hz for 10 seconds) for each of the 12 leads and outputs one singular 512-dimensional feature vector. In the first approach, the ResNet-18 was initialized for hospitalization prediction with random weights. In the second, it was first trained on a larger ECG dataset with labels unrelated to ED disposition, to capture general waveform structure and patterns. Training beforehand allows the network to initialize with physiologically relevant weights, enhancing feature extraction of QRS-T morphology and inter-lead patterns, while significantly reducing overfitting through faster convergence and freedom to freeze layers. Reduced overfitting may be important given the added complexity of the multimodal architecture.

As shown in Figure 2, the full MIMIC-IV-ECG dataset was used as a means of supervised pretraining, in which each ECG waveform was used as an input to predict MIMIC machine measurements and machine-generated reports. These measurements and reports are provided on the MIMIC-IV-ECG module for every available ECG. Machine measurements reflect quantitative ECG characteristics across all leads, including the average RR interval, QRS axis, T-axis, and several more. Machine-generated reports are stored as strings in columns labeled report 1 through report 17, with example strings being "Atrial Fibrillation," "ST-elevation," or "Normal ECG." After one-hot encoding, 3129 columns were initially created. After removing labels that had too few positive instances, along with combining truth values of columns with slight syntax variance or similar clinical significance, such as "ventricular pacing" and "ventricular-paced rhythm," 82 unique Boolean features remained.

**Figure 2.** Flowchart of supervised pretraining workflow for the transfer-learning approach. D: dimensional; ECG: electrocardiogram; MIMIC-IV-ECG: Medical Information Mart for Intensive Care IV electrocardiogram module; ResNet-18: residual network model.




Before training, all ECGs were cleaned using the Neurokit2 signal processing library [31]. A one-dimensional ResNet-18 model was trained to predict both numeric and Boolean features corresponding to each of the total 7,57,754 ECGs. Mean squared error was used to evaluate the loss of numeric features (machine measurements), and binary cross-entropy (BCE) was used to evaluate the loss of Boolean features (machine-generated reports). The total loss function represented the mean squared error loss added to the BCE loss multiplied by 9 (due to far more Boolean labels than numeric).

## Prediction Model

### *Multimodal*

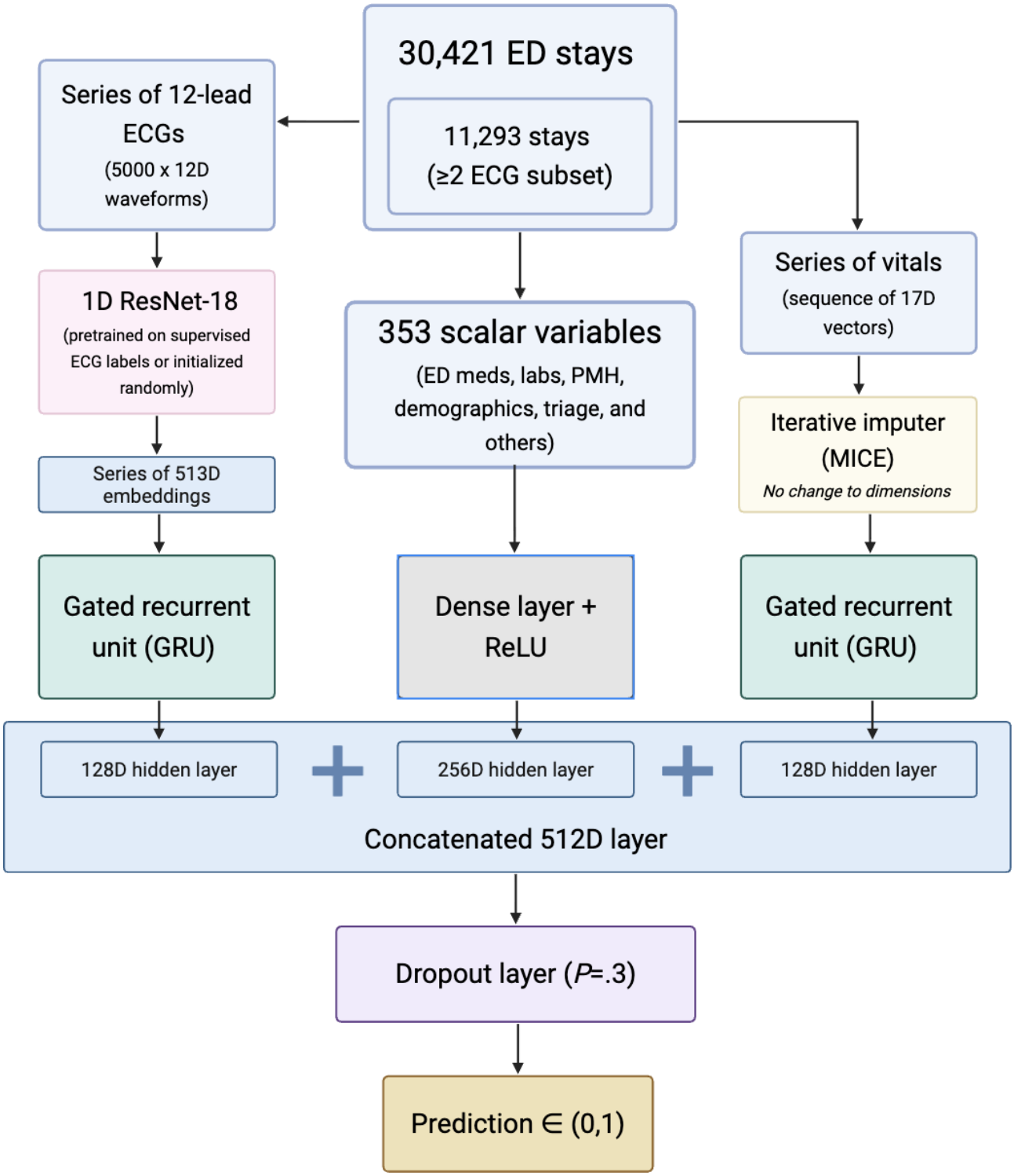
A multimodal model trained on sequential ECG waveforms, sequential vitals data, and 353 static variables was built to predict hospitalization in patients presenting to the ED. The model used all data gathered during the stay before the final ECG. The model was trained on all stays as well as on a subset of stays containing at least 2 ECGs to quantify the importance of sequential ECGs and more available data within the time window.

Raw 12-lead ECGs were first cleaned with NeuroKit2 and subsequently one-dimensional ResNet-18 initialized with random weights or the ResNet-18 that had been trained on the larger MIMIC-IV-ECG dataset (757,754 recordings). In the transfer-learning approach, ResNet-18 layers through block layer3 were frozen: only layer4 and the linear adapter were trained during finetuning. The final global-average-pooling layer of the backbone was replaced with a linear adapter so that each waveform was mapped to a fixed-length, 512-dimensional embedding.

For every ED stay we retained, in temporal order, all ECGs recorded before the clinical disposition decision (maximum 6 per stay) and all vital sign rows charted before the last ECG (maximum 10 per stay). The ECG and vital embeddings were first stacked and padded to match the maximum length. As shown in Figure 3, 2 independent, single-layer gated recurrent units (GRUs) were used to summarize these sequences. Before being passed to the GRU, padded batches were converted to PackedSequence objects, ensuring the recurrent unit was unrolled only over the valid timesteps and ignored the artificial zero-padding. The time before the final ECG was added as an extra time-delta channel in the embeddings of each ECG and

vitals sequence. The ECG GRU read the sequence of 513-dimensional embeddings and returned a 128-dimensional hidden state $h_{ECG}$; the vitals GRU processed a 17-variable vector at each timestep—7 z-scored physiological values, 1 time-delta channel, and 9 binary mask indicators denoting whether the original measurement had been missing, as well as unable or sleeping for "pain"—and produced a similar 128-dimensional summary $h_{VITALS}$.

**Figure 3.** Diagram of the multimodal model architecture approach to predict hospitalization in the full and subset cohorts. D: dimensional; ED: emergency department; PMH: past medical history; ReLU: rectified linear unit; ResNet: residual network model.




All 353 static features were concatenated, normalized with the mean (SD) calculated on the training set, and projected by a fully connected layer with rectified linear unit activation to a 256-dimensional vector *s*.

The 3 modality-specific representations were then fused through simple concatenation,

$z = [\, h_{ECG} \| h_{VITALS} \| s \,]^{512}$, followed by an additional dropout layer (rate=0.30) and a single neuron that yielded the estimated probability of hospital admission $\hat{p} \in (0,1)$.

Model parameters introduced during this study were optimized with AdamW (weight decay $10^{-4}$ at a learning rate of $3 \times 10^{-4}$ [32]; the adapter atop the ECG encoder was updated at $10^{-4}$. BCE with class weighting was used as the optimization objective.

### *Baseline*

To measure the effect of ECGs alone in predicting disposition, separate models were trained on only ECG features. The ECG-only model used the same parameters and architecture as the multimodal ECG branch.

Random forest (RF) is an ML algorithm that combines predictions from an ensemble of decision trees to produce 1 final classification. An RF was constructed solely on the 353 static tabular features as a baseline to the multimodal architecture. Both the ECG-only and tabular baseline models were evaluated in the full cohort ("all stays") and in the ≥2-ECG subset using the exact stratified 5-fold cross-validation splits from the multimodal pipeline.

### *Model Evaluation*

Three models (multimodal, tabular, and ECG-only) were evaluated on the full cohort of all stays and on the ≥2-ECG subset. Across cross-validation folds, the mean AUROC, area under the precision-recall curve, $F_1$-score, precision, and recall were reported. Thresholds for $F_1$-score, precision, and recall were selected using a nested procedure that maximized $F_1$-score on training folds and applied it to the validation fold. Pairwise AUROC differences between models were tested with the DeLong method.

### Ethical Considerations

This study used deidentified data from the PhysioNet MIMIC-IV database under a data use agreement. Because the data are fully deidentified, the study was determined to be exempt from institutional review board (IRB) review, and informed consent was waived in accordance with applicable regulations. All data handling complied with relevant privacy and data protection standards, and no identifiable personal information was accessed. No participant contact or compensation occurred. This research was conducted in accordance with the ethical standards of the institutional and national research committees and with the 1975 Declaration of Helsinki (as revised in 2000).

## Results

### ED Timing Analysis

Of the 30,421 ED stays included in our analysis, 11,273 (37.1%) involved stays with ≥2 ECGs (Table 1). In the full cohort, the model issued its prediction a median of 0.3 (IQR 0.2 - 5.3) hours after triage, leaving a median of 4.6 (IQR 2.7 - 7.3) hours before the patient physically left the ED. Note that these medians are calculated independently and therefore do not sum to the overall median length of stay. As a result, disposition was predicted in the first 18 minutes for most visits, but for the upper quartile, the model leveraged data gathered more than 5 hours into the encounter. In the ≥2 ECG subset, the model predicted disposition far later, with a median of 6.5 (IQR 4.3 - 9.8) hours after triage. The median time from prediction to disposition was also shorter than the full cohort, at 3.4 (IQR 1.7 - 9.1) hours. However, the 75th-percentile (Q3) prediction lead time exceeded 9 hours, meaning the model anticipated disposition more than 9 hours before the actual decision. This subset also had a longer overall ED length of stay (median 9.5, IQR 6.9 - 17.8 hours vs 6.5, IQR 4.3 - 9.8 hours in the full cohort). Because the ≥2 ECG cohort had more ECGs taken and the model predicts disposition after the final ECG, it often took more time to reach the final ECG.

**Table .** Baseline characteristics and measurements in all versus ≥2-ECG[a] cohorts.

| Metrics[b] | All stays (N=30,421) | ≥2 ECG stays (n=11,273) |
|---|---|---|
| Admit, n (%) | 13,138 (43.2) | 4070 (36.1) |
| Acuity, mean (SD) | 2.3 (0.6) | 2.2 (0.6) |
| Age (years), mean (SD) | 57.1 (19.8) | 61.7 (16.3) |
| ECG metrics | | |
| Number of ECGs, mean (SD) | 1.48 (0.74) | 2.30 (0.64) |
| ≥2 ECGs, n (%) | 11,273 (37.1) | 11,273 (100) |
| ≥3 ECGs, n (%) | 2538 (8.3) | 2538 (22.5) |
| LOS[c] after final ECG (hours), median (IQR) | 4.6 (2.7 - 7.3) | 3.4 (1.7 - 9.1) |
| LOS before final ECG (hours), median (IQR) | 0.3 (0.2 - 5.3) | 6.5 (3.9 - 7.7) |
| Full LOS (hours), median (IQR) | 6.5 (4.3 - 9.8) | 9.5 (6.9 - 17.8) |
| Laboratory values (first) | | |
| Troponin T (ng/mL), median (IQR) | 0.1 (0 - 0.2) | 0.1 (0 - 0.2) |
| Missing, n (%) | 28,376 (93.3) | 9579 (85) |
| Lactate (mmol/L), median (IQR) | 1.7 (1.3 - 2.3) | 1.7 (1.3 - 2.3) |
| Missing, n (%) | 27,856 (91.6) | 9206 (81.7) |
| Creatinine (mg/dL), median (IQR) | 0.9 (0.8 - 1.1) | 0.9 (0.8 - 1.1) |
| Missing, n (%) | 18,327 (60.2) | 1177 (10.4) |
| B-type natriuretic peptide (pg/mL), median (IQR) | 1092.5 (180 - 4555.8) | 948.5 (162 - 4293.2) |
| Missing, n (%) | 27,993 (92) | 9363 (83.1) |
| Potassium (mmol/L), median (IQR) | 4.2 (3.9 - 4.6) | 4.2 (3.9 - 4.6) |
| Missing, n (%) | 18,340 (60.3) | 1170 (10.4) |
| Magnesium (mg/dL), median (IQR) | 2 (1.9 - 2.2) | 2 (1.9 - 2.2) |
| Missing, n (%) | 26,702 (87.8) | 8244 (73.1) |
| Presenting symptoms, n (%) | | |
| Chest pain | 17,898 (58.8) | 8519 (75.6) |
| Dyspnea | 9477 (31.2) | 2575 (22.8) |
| Syncope | 4924 (16.2) | 983 (8.7) |
| Summary counts, n (%) | | |
| No visible ED[d] medications before final ECG | 21,368 (70.2) | 3583 (31.8) |
| No visible prior medications (past year) | 22,074 (72.6) | 8157 (72.4) |
| No visible prior hospital diagnosis | 15,950 (52.4) | 5275 (46.8) |
| No visible prior ED stay | 16,594 (54.5) | 5591 (49.6) |

[a]ECG: electrocardiogram.

[b]Percentages for all rows are calculated using unique patients as the denominator (N). The values from the selected representative lab measurements are from tested patients only.

[c]LOS: length of stay.

[d]ED: emergency department.

Detailed vitals summaries, including missingness counts by cohort, are shown in Table S1 in Multimedia Appendix 1.

## Missing Values

When vitals were charted, only temperature was frequently missing within 41.8% (28,573/68,371) of vital chartings within all stays and 55% (25,376/46,098) in the ≥2 ECG subset (Table S1 in Multimedia Appendix 1). All other vitals were <7% missing per charting, and many stays had multiple vitals chartings before the prediction cutoff. Of the 9 lab values this study accounted for, 6 were entirely missing in more than 90% (~27,400/30,421) of all patients before the prediction cutoff

(Table 1). Because the subset included a longer median interval before the prediction cutoff, the proportion of missing laboratory results fell for every test. Notably, creatinine dropped from 60.2% (18,327/30,421) to 10.4% (1177/11,273), missing between the cohorts. PMH within the dataset's visibility was very limited. Any prior diagnoses and ED stays were each completely missing in more than half of the total patients. A total of 72.6% (22,074/30,421) of patients had no visible medications prescribed to them in the past year in the dataset.

## ECG Feature Extractor

Compared with random initialization, supervised transfer learning of the ResNet-18 did not yield statistically significant improvements in predictive performance. However, supervised pretrained models consistently converged faster and achieved marginally higher metrics in the ≥2 ECG subset. Accordingly, all reported results used the transfer-learning approach.

## Model Evaluation

Using a partially frozen pretrained ResNet-18 as an ECG encoder, a multimodal dual GRU fusion net was trained on the full cohort and ≥2 ECG subset to predict hospital admission. For comparison, an ECG-only variant (same encoder and GRU using only raw 12-lead waveforms) and a tabular-only RF were trained for the same task. Performance is shown in Table 2.

**Table .** Performance metrics for ECG[a]-only, multimodal, and tabular random forest models. Evaluations use an identical time cutoff across models and are stratified by cohort (all stays; ≥2 ECGs). Expected calibration error was computed from 10 quantile bins. *P* values are from paired DeLong tests versus the multimodal model. Metrics are mean values across the 5-fold cross-validation.

| Models | AUROC[b], mean (SD) | AUPRC[c], mean (SD) | Precision, mean (SD) | Recall, mean (SD) | ECE[d], mean (SD) | *P* value |
|---|---|---|---|---|---|---|
| All stays | | | | | | |
| ECG only | 0.852 (0.003) | 0.813 (0.005) | 0.698 (0.034) | 0.803 (0.036) | 0.034 (0.018) | <.001 |
| Tabular (RF[e]) | 0.886 (0.003) | 0.849 (0.006) | 0.745 (0.017) | 0.830 (0.014) | 0.024 (0.004) | <.001 |
| Multimodal | 0.911 (0.004) | 0.889 (0.005) | 0.784 (0.041) | 0.839 (0.046) | 0.026 (0.008) | —[f] |
| ≥2 ECG stays | | | | | | |
| ECG only | 0.859 (0.011) | 0.794 (0.017) | 0.674 (0.029) | 0.760 (0.038) | 0.053 (0.018) | <.001 |
| Tabular (RF) | 0.911 (0.006) | 0.865 (0.014) | 0.774 (0.016) | 0.813 (0.013) | 0.039 (0.008) | <.001 |
| Multimodal | 0.924 (0.009) | 0.889 (0.016) | 0.807 (0.030) | 0.808 (0.024) | 0.040 (0.024) | — |

[a]ECG: electrocardiogram.

[b]AUROC: area under receiver operating characteristic.

[c]AUPRC: area under the precision-recall curve.

[d]ECE: expected calibration error.

[e]RF: random forest.

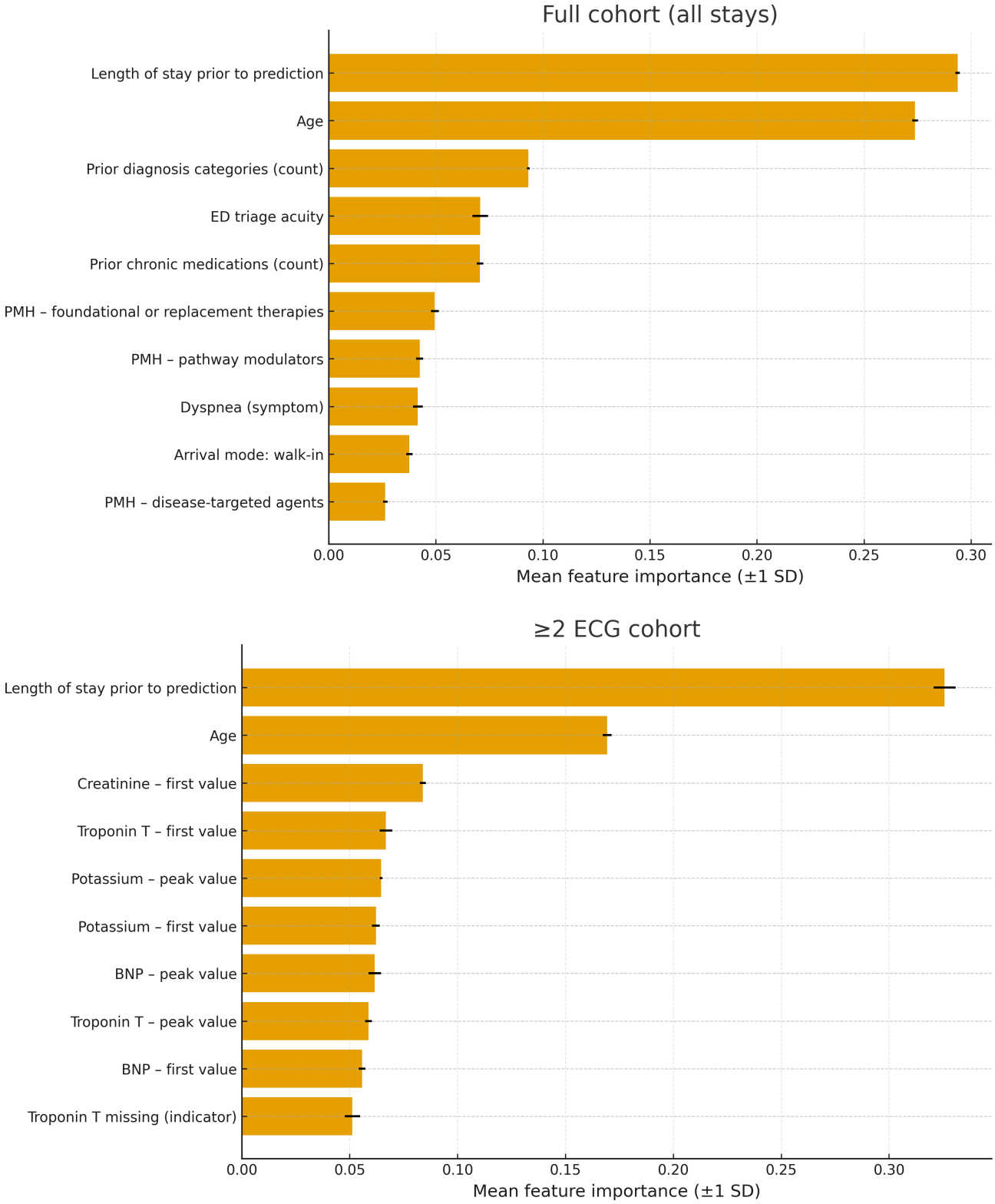
[f]Not applicable.

The multimodal model in the multiple ECG subset yielded the highest metrics, with an AUROC of 0.924. The tabular RF model yielded slightly lower metrics (AUROC=0.91). Across cohorts, the AUROC of the ECG-only model differed by as little as AUROC of 0.007, suggesting that additional information from prior ECGs in the same stay does not improve prediction by a significant margin. In the all-stays cohort, where predictions were made much earlier in the ED course, the performance gap between the random forest and multimodal models was larger: AUROC 0.886 versus 0.911 ($P$<.001). These findings indicate that combining ECG waveforms with vitals provides the largest benefit when static tabular data, such as lab results, are sparse early in the encounter.

## Static Feature Importance

The 10 static features with the highest mean feature importance in both cohorts for the tabular RF model are displayed in Figure 4. All 4 features corresponding to Troponin were included in the top 10 in the multiple ECG subset, but none appear in the full cohort of all stays. The greater importance of lab values in the multiple ECG subset is most likely due to less missing lab data, as the median cutoff for prediction later. The length of stay prior to prediction and the patient's age remain the most important features within both cohorts. In the full cohort, PMH and acuity were given more importance.

**Figure 4.** Two bar graphs showing mean feature importance across 5 folds for both cohorts. BNP: B-type natriuretic peptide; ECG: electrocardiogram; ED: emergency department; PMH: past medical history.




We also conducted an ablation study evaluating a tabular RF trained on only the top 10 static features (vs all 353). On the multi-ECG subset, the reduced model achieved an AUROC of 0.862 compared to 0.911 with the full feature set. On the full cohort, performance decreased from 0.886 to 0.841 when limited to 10 features.

## Discussion

### Principal Findings

In this study, we developed a multimodal deep-learning model that integrates sequential ECG waveforms and vital signs, along with static tabular features, to predict hospitalization in real time for patients presenting to the ED with possible cardiac

presentations. Our results show that ECGs can also be used alone to predict disposition, but are best used in combination with vitals and other commonly implemented static features such as lab results, medications administered, prior medical history, and triage data. The primary operational value of multimodal fusion appears early when tabular data are sparse, whereas later in the visit, a simpler tabular model may achieve comparable discrimination at lower implementation cost.

The multimodal model that required at least 2 ECGs achieved the highest performance of AUROC of 0.924 in predicting hospitalization. This model predicted a disposition of 3.4 (IQR 1.7 - 9.1) hours in advance on the median, but more than 9 (median 9.5, IQR 6.9 - 17.8) hours in advance for the top 25th percentile with high accuracy. However, a median of 6.5 (IQR 3.9 - 7.7) hours was spent in the ED before this model reached a prediction. The RF built from tabular data achieved a slightly lower AUROC of 0.911. As the ECG-only model's performance did not improve significantly with more ECGs, the higher AUROC in the ≥2 ECG subset for the multimodal and RF models is most likely due to the more lenient cutoff to chart lab results and ED medications before prediction.

When trained on all-stays (median prediction time less than 20 minutes after triage), the multimodal architecture yielded significantly higher metrics in comparison to tabular and ECG-only models. Incorporating additional data streams such as waveforms and vitals can lead to better and more reliable hospitalization predictions when traditional clinical data are sparse.

## Comparison With Prior Work

Our ECG-based prediction models substantially outperform conventional triage tools such as the nurse-assigned ESI level, which typically achieves AUROCs in the 0.69 - 0.70 range and is at par with or better than many prior ML models that used only triage-time data [11]. For example, Raita et al [11] reported an AUROC of 0.82 for a deep neural network predicting hospitalization using initial-only triage information, and Hong et al [12] similarly found approximately 0.87 using triage data with extreme gradient boosting and a neural network. In addition, Hong et al [12] also created extreme gradient boosting and deep neural network models, which reached an AUROC of 0.92, implementing PMH and triage data to predict hospital admission at the beginning of the ED stay. Their approach is further discussed in the limitations section of this study.

Some studies make use of data after triage to predict disposition. Barak-Corren et al [24] achieved high performance (AUROC=0.97 within 1 h of triage) using logistic regression on standard demographic, triage, medication, lab, vital, and PMH data in a single-site Israeli ED. However, they train their logistic regression model to predict disposition on a per-visit basis, while we provide a per-patient approach, a more challenging task because there is no patient overlap within the train and validation cohort. Sezik et al [22] used similar data (lab results, history, and triage variables) in combination with vitals and other features extracted from text to reach an AUROC of 0.960 with an RF classifier. However, they do not incorporate a cutoff time and use features conducted throughout the entire ED stay, including the length of stay. In contrast, our study stops feature collection after a given ECG in order to predict later changes in a patient presenting with a cardiac condition and provide an early hospitalization prediction as opposed to serving as an aid to decide disposition at the end of the stay.

Importantly, our approach maintained high accuracy despite the broad outcome of general hospital admission, which includes many noncritical cases; this is notable because prior ML studies in chest pain often focused on narrower critical events. For instance, a recent study targeted only critical care outcomes such as ED cardiac arrest or intensive care unit transfer and achieved an AUROC of ~0.95 with a tailored Least Absolute Shrinkage and Selection Operator logistic model [13]. The fact that we still attained strong performance (AUROC=0.911 and 0.924) for a broader outcome of hospitalization for any reason underlines the effectiveness of our multimodal fusion strategy. Using the MIMIC-IV-ED and MIMIC-IV-ECG modules (which include more than 400,000 ED stays and 800,000 ECGs, respectively) ensured that our work is widely reproducible [29,30]. This reproducibility and open-source approach contrasts with many prior ED prediction studies that relied on proprietary data not available to outside researchers [9,12-15,17-24]. By creating our model from a public database, we demonstrate the feasibility of developing advanced prediction tools using open data.

## Limitations and Future Work

This study has several limitations. First, all data used to train the model are specific to the single-center MIMIC-IV database from the Beth Israel Deaconess Medical Center. This limits generalizability, as the model may overfit to site-specific patterns in the high-dimensional input space. Thus, external validation is needed. Nonetheless, much of the data used, such as common 12-lead ECGs, vitals, and standardized eHealth records, could be readily translated in future multicenter studies to support broader applicability.

Second, due to exploring the promise of ECG waveforms on ED disposition, our study relies on the patient presenting with a possible cardiac condition, which in this case only includes patients presenting with chest pain, dyspnea, presyncope, and syncope. These symptoms were chosen as they are the most common in the ED and are predominantly cardiac in origin, in contrast to less frequent and more heterogeneous complaints such as palpitations. In addition, all included patients received at least 1 ECG, and although it is standard for patients with possible cardiac symptoms to promptly receive an ECG, this could still affect the model's generalizability. Future work should test whether expanding the cohort further broadens the model's applicability.

Third, many variables were largely missing, including more than 90% of most labs tested and 41% to 55% of temperature readings when vitals were charted. As opposed to the 0.92 AUROC neural network proposed by Hong et al [12], which contained full patient medical records and a comprehensive PMH of each patient, we only included past history visible within the MIMIC-IV and MIMIC-IV-ED modules. Overall, 72.6% (22,074/30,421) of patients in our study set had no prior medications, 52.4% (15,950/30,421) had no prior diagnoses, and 54.5% (16,594/30,421) had no prior ED stays. PMH

recorded at another hospital was not visible. Considering the value shown by incorporating a comprehensive PMH in predicting disposition, our model performance could significantly benefit from having access to more prior diagnoses, medications, and ED stays.

Finally, like most disposition models, our model treats the disposition recommended by the ED provider as a truth label. Thus, our model may be, to some extent, learning institutional decision policies that can vary across hospitals in addition to patient physiology. The fact that "Length of stay prior to prediction" was the most predictive static variable likely reflects a confounding effect. Longer ED stays often indicate physician uncertainty or patient complexity, meaning the model may capture care processes as well as outcomes. Although discrimination was highest in the ≥2-ECG subset, its prediction was issued a median of 6.5 (IQR 4.3 - 9.8) hours after triage (~3.4 hours before ED departure), which may limit clinical use for throughput and flow optimization.

In practice, the model would run automatically after each ECG, showing a calibrated admission probability and risk band in the EHR as a banner alert. The model could also be rerun with new vitals or lab measurements, which are automatically entered into the patient's record. At inference, the system runs only forward passes through a 1D ResNet-18 encoder and small GRUs, making it suitable for near-real-time usage. We envision it as decision support that complements existing scores early in the visit, with a tiered strategy where a simpler tabular model can suffice later, once more labs and medications are available. Future work should explore including more data modalities in a time-series format and possibly leaving the cutoff for prediction variable throughout updates in lab testing, vitals, and diagnostic tools such as an ECG. Incorporating novel data streams, such as the ECG with a more complete PMH, could lead to a far better prediction as well. Improving model interpretability, especially within deep-learning models, is also crucial to gain clinician trust and improve decision-making in the ED.

## Conclusions

We developed and validated a multimodal deep learning model that combines ECG waveforms, vital signs, and static clinical features to predict hospital admission in patients presenting to the ED with potential cardiac complaints. The model achieved a higher AUROC than both ECG-only and tabular-only baselines and issued predictions at the time of the final ECG, a median of 0.3 (IQR 0.2 - 5.3) hours after triage. These findings suggest that integrating ECG waveforms with sequential and static data may enhance early risk stratification in the tested patient population. Future external validation is required to assess generalizability.

## Acknowledgments

The authors thank Ashwath Radhachandran and Amir Reza Vazifeh for providing valuable feedback on the article. This research received no external funding.

## Data Availability

The study draws exclusively on the publicly available Medical Information Mart for Intensive Care version IV (MIMIC-IV) databases—including the emergency-department (MIMIC-IV-ED) and ECG (MIMIC-IV-ECG) modules—hosted on PhysioNet. Because access to MIMIC-IV is restricted under a data-use agreement that protects patient privacy, raw or preprocessed patient-level data cannot be redistributed here. Interested researchers can obtain the same data free of charge by completing the data-use certification on PhysioNet. All scripts for data preprocessing, model training, and evaluation, as well as trained model weights, are openly available at GitHub [33].

## Authors' Contributions

AA: conceptualization, methodology, data curation, formal analysis, investigation, software, validation, visualization, writing–original draft, and writing–review and editing. KBO: supervision, validation, and writing–review and editing.

## Conflicts of Interest

None declared.

Multimedia Appendix 1
Vitals measurements in all versus ≥2 electrocardiogram cohorts.
[DOCX File, 14 KB - cardio_v9i1e80569_app1.docx ]

## References

1. Gupta M, Tabas JA, Kohn MA. Presenting complaint among patients with myocardial infarction who present to an urban, public hospital emergency department. Ann Emerg Med 2002 Aug;40(2):180-186. [doi: 10.1067/mem.2002.126396] [Medline: 12140497]
2. Yukselen Z, Majmundar V, Dasari M, Arun Kumar P, Singh Y. Chest pain risk stratification in the emergency department: current perspectives. Open Access Emerg Med 2024;16:29-43. [doi: 10.2147/OAEM.S419657] [Medline: 38343728]

3. Patel PR, Quinn JV. Syncope: a review of emergency department management and disposition. Clin Exp Emerg Med 2015 Jun;2(2):67-74. [doi: 10.15441/ceem.14.049] [Medline: 27752576]
4. DeVos E, Jacobson L. Approach to adult patients with acute dyspnea. Emerg Med Clin North Am 2016 Feb;34(1):129-149. [doi: 10.1016/j.emc.2015.08.008] [Medline: 26614245]
5. McDevitt-Petrovic O, Kirby K, Shevlin M. The prevalence of non-cardiac chest pain (NCCP) using emergency department (ED) data: a Northern Ireland based study. BMC Health Serv Res 2017 Aug 9;17(1):549. [doi: 10.1186/s12913-017-2493-8] [Medline: 28793910]
6. Maisel AS, Krishnaswamy P, Nowak RM, et al. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. N Engl J Med 2002 Jul 18;347(3):161-167. [doi: 10.1056/NEJMoa020233] [Medline: 12124404]
7. Collins SP, Jenkins CA, Harrell FE Jr, et al. Identification of emergency department patients with acute heart failure at low risk for 30-day adverse events: the STRATIFY decision tool. JACC Heart Fail 2015 Oct;3(10):737-747. [doi: 10.1016/j.jchf.2015.05.007] [Medline: 26449993]
8. Hatoum T, Sheldon RS. A rational evaluation of the syncope patient: optimizing the emergency department visit. Medicina (Kaunas) 2021 May 21;57(6):514. [doi: 10.3390/medicina57060514] [Medline: 34064050]
9. Levin S, Toerper M, Hamrock E, et al. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. Ann Emerg Med 2018 May;71(5):565-574. [doi: 10.1016/j.annemergmed.2017.08.005] [Medline: 28888332]
10. Bernstein SL, Aronsky D, Duseja R, et al. The effect of emergency department crowding on clinically oriented outcomes. Acad Emerg Med 2009 Jan;16(1):1-10. [doi: 10.1111/j.1553-2712.2008.00295.x] [Medline: 19007346]
11. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care 2019 Feb 22;23(1):64. [doi: 10.1186/s13054-019-2351-7] [Medline: 30795786]
12. Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. PLoS One 2018;13(7):e0201016. [doi: 10.1371/journal.pone.0201016] [Medline: 30028888]
13. Wu TT, Zheng RF, Lin ZZ, Gong HR, Li H. A machine learning model to predict critical care outcomes in patient with chest pain visiting the emergency department. BMC Emerg Med 2021 Oct 7;21(1):112. [doi: 10.1186/s12873-021-00501-8] [Medline: 34620086]
14. Gustafsson S, Gedon D, Lampa E, et al. Development and validation of deep learning ECG-based prediction of myocardial infarction in emergency department patients. Sci Rep 2022 Nov 15;12(1):19615. [doi: 10.1038/s41598-022-24254-x] [Medline: 36380048]
15. Peck JS, Benneyan JC, Nightingale DJ, Gaehde SA. Predicting emergency department inpatient admissions to improve same-day patient flow. Acad Emerg Med 2012 Sep;19(9):E1045-E1054. [doi: 10.1111/j.1553-2712.2012.01435.x] [Medline: 22978731]
16. Dugas AF, Kirsch TD, Toerper M, et al. An electronic emergency triage system to improve patient distribution by critical outcomes. J Emerg Med 2016 Jun;50(6):910-918. [doi: 10.1016/j.jemermed.2016.02.026] [Medline: 27133736]
17. Leegon J, Jones I, Lanaghan K, Aronsky D. Predicting hospital admission in a pediatric emergency department using an artificial neural network. AMIA Annu Symp Proc 2006;2006:1004. [Medline: 17238623]
18. Leegon J, Jones I, Lanaghan K, Aronsky D. Predicting hospital admission for emergency department patients using a bayesian network. AMIA Annu Symp Proc 2005;2005:1022. [Medline: 16779309]
19. Lucini FR, Fogliatto FS, da Silveira GJC, et al. Text mining approach to predict hospital admissions using early medical records from the emergency department. Int J Med Inform 2017 Apr;100:1-8. [doi: 10.1016/j.ijmedinf.2017.01.001] [Medline: 28241931]
20. Lucke JA, de Gelder J, Clarijs F, et al. Early prediction of hospital admission for emergency department patients: a comparison between patients younger or older than 70 years. Emerg Med J 2018 Jan;35(1):18-27. [doi: 10.1136/emermed-2016-205846] [Medline: 28814479]
21. Chang YH, Lin YC, Huang FW, et al. Using machine learning and natural language processing in triage for prediction of clinical disposition in the emergency department. BMC Emerg Med 2024 Dec 18;24(1):237. [doi: 10.1186/s12873-024-01152-1] [Medline: 39695961]
22. Sezik S, Cingiz M, İbiş E. Machine learning-based model for emergency department disposition at a public hospital. Appl Sci (Basel) 2025;15(3):1628. [doi: 10.3390/app15031628]
23. Barak-Corren Y, Israelit SH, Reis BY. Progressive prediction of hospitalisation in the emergency department: uncovering hidden patterns to improve patient flow. Emerg Med J 2017 May;34(5):308-314. [doi: 10.1136/emermed-2014-203819] [Medline: 28188202]
24. Barak-Corren Y, Fine AM, Reis BY. Early prediction model of patient hospitalization from the pediatric emergency department. Pediatrics 2017 May;139(5):e20162785. [doi: 10.1542/peds.2016-2785] [Medline: 28557729]
25. Bouzid Z, Faramand Z, Martin-Gill C, et al. Incorporation of serial 12-lead electrocardiogram with machine learning to augment the out-of-hospital diagnosis of non-ST elevation acute coronary syndrome. Ann Emerg Med 2023 Jan;81(1):57-69. [doi: 10.1016/j.annemergmed.2022.08.005] [Medline: 36253296]
26. Johnson AEW, Bulgarelli L, Pollard TJ, et al. MIMIC-IV (version 3.0). PhysioNet 2024. [doi: 10.13026/hxp0-hg59]

27. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data 2023 Jan 3;10(1):1. [doi: 10.1038/s41597-022-01899-x] [Medline: 36596836]
28. Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 2000 Jun 13;101(23):E215-E220. [doi: 10.1161/01.cir.101.23.e215] [Medline: 10851218]
29. Johnson AEW, Bulgarelli L, Pollard TJ, Celi LA, Mark RG, Horng S. MIE. MIMIC-IV-ED (version 2.2). PhysioNet 2023. [doi: 10.13026/5ntk-km72]
30. Gow B, Pollard T, Nathanson LA, et al. MIMIC-IV-ECG: diagnostic electrocardiogram matched subset (version 1.0). PhysioNet 2023. [doi: 10.13026/4nqg-sb35]
31. Makowski D, Pham T, Lau ZJ, et al. NeuroKit2: a python toolbox for neurophysiological signal processing. Behav Res Methods 2021 Aug;53(4):1689-1696. [doi: 10.3758/s13428-020-01516-y] [Medline: 33528817]
32. Loshchilov I, Hutter F. Decoupled weight decay regularization. Presented at: ICLR 2019-The Seventh International Conference on Learning Representations; May 6-9, 2019; New Orleans, LA URL: https://arxiv.org/abs/1711.05101 [accessed 2025-10-08]
33. Altintepe A. Serial 12-lead ECG–based deep-learning model for real-time prediction of hospital admission in emergency-department cardiac presentations: retrospective cohort study. GitHub. URL: https://github.com/againh3x/ECGhospitalization [accessed 2025-10-08]


## Abbreviations

**AUROC:** area under the receiver operating characteristic curve
**BCE:** binary cross-entropy
**ED:** emergency department
**ESI:** emergency severity index
**GRU:** gated recurrent unit
**HEART:** History, ECG, Age, Risk factors, and Troponin
**MIMIC-IV:** Medical Information Mart for Intensive Care IV
**MIMIC-IV-ECG:** MIMIC-IV electrocardiogram module
**MIMIC-IV-ED:** MIMIC-IV Emergency Department module
**ML:** machine learning
**PMH:** past medical history
**ResNet-18:** residual neural network
**RF:** random forest


*Edited by A Coristine; submitted 25.08.25; peer-reviewed by G Thomas Brown, KH Lin; revised version received 29.09.25; accepted 30.09.25; published 17.10.25.*

*Please cite as:*
*Altintepe A, Ozyoruk KB*
*Serial 12-Lead Electrocardiogram–Based Deep-Learning Model for Hospital Admission Prediction in Emergency Department Cardiac Presentations: Retrospective Cohort Study*
*JMIR Cardio 2025;9:e80569*
*URL: https://cardio.jmir.org/2025/1/e80569*
*doi:10.2196/80569*

# Gender Differences in X (Formerly Twitter) Use, Influence, and Engagement Among Cardiologists From the Top U.S. News Best Hospitals

Minji Seok[1], MD; Sungjin Kim[2], MS; Harper Tzou[3]; Olivia Peony[3]; Mitchell Kamrava[3], MSc, MD; Andriana P Nikolova[4], MD, PhD; Katelyn M Atkins[3,4], MD, PhD

[1]Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, United States

[2]Biostatistics Research Center, Cedars-Sinai Medical Center, Los Angeles, CA, United States

[3]Department of Radiation Oncology, Cedars-Sinai Medical Center, 8700 Beverly Blvd AC1019, Los Angeles, CA, United States

[4]Department of Cardiology, Cedars-Sinai Medical Center, Los Angeles, CA, United States

**Corresponding Author:**
Katelyn M Atkins, MD, PhD
Department of Radiation Oncology, Cedars-Sinai Medical Center, 8700 Beverly Blvd AC1019, Los Angeles, CA, United States

***(JMIR Cardio 2025;9:e66308)*** doi:10.2196/66308

**KEYWORDS**

social media; women in medicine; health communication; technology; Twitter; cardiologist; cardiology; US; United States; heart; vascular surgery; adult; women; female; medicine

## Introduction

Women in medicine face significant barriers to compensation, career advancement, and research support, even when controlling for specialty, age, and/or clinical experience [1]. These barriers are especially pronounced in cardiology, where women comprise only 15% of practicing cardiologists and are less likely to be clinical trial leaders or present late-breaking trials at major cardiovascular conferences [2-4]. Social media platforms, such as X (formerly Twitter), can foster collaboration, mentorship, and promotion of research [5,6]. However, studies examining X's impact on existing gender gaps are limited. In this study, we aimed to analyze differences between X users and non–X users and differences in X use by gender among adult cardiologists.

## Methods

### Ethical Considerations

This cross-sectional study was exempt from ethical approval by the Cedars-Sinai institutional review board due to the use of publicly available data.

### Study Design

The top 20 U.S. News Best Hospitals for cardiology, heart surgery, and vascular surgery were identified from the 2023 ranking (Table 1) [7]. Available physician website profiles of fellowship-trained adult medicine cardiologists were manually reviewed by 3 investigators (MS, HT, and OP) for inclusion, and demographic information was collected (eg, academic appointment, apparent gender, and medical school and fellowship graduation years). Physicians were evaluated for the presence of an X account, and public data were manually collected between December 8, 2023, and May 9, 2024. Differences between non–X users and X users and between women and men X users were compared, using Wilcoxon rank-sum tests for continuous variables and chi-square or Fisher exact tests for categorical variables as appropriate.

**Table .** Top 20 U.S. News Best Hospitals for cardiology, heart surgery, and vascular surgery (2023 ranking).

| Institution name | State | Total physicians (N=2022), n (%) | Physicians on X (n=753), n (%) |
|---|---|---|---|
| Brigham and Womens | Massachusetts | 143 (7.07) | 73 (9.69) |
| Cedars Sinai | California | 56 (2.77) | 22 (2.92) |
| Cleveland Clinic | Florida and Ohio | 126 (6.23) | 51 (6.77) |
| Johns Hopkins | Maryland | 102 (5.04) | 35 (4.65) |
| Houston Methodist | Texas | 64 (3.17) | 33 (4.38) |
| Lenox Hill at Northwell | New York | 117 (5.79) | 27 (3.59) |
| Massachusetts General | Massachusetts | 100 (4.95) | 57 (7.57) |
| Mayo Clinic Rochester | Minnesota | 156 (7.72) | 68 (9.03) |
| Mount Sinai | Florida, New Jersey, and New York | 201 (9.94) | 64 (8.50) |
| NewYork-Presbyterian Hospital Columbia and Cornell | New York | 54 (2.67) | 15 (1.99) |
| NYU Langone Hospitals | New York | 164 (8.11) | 20 (2.66) |
| Northwell Northshore | New York | 93 (4.60) | 15 (1.99) |
| Northwestern | Illinois | 112 (5.54) | 47 (6.24) |
| Rush University | Illinois | 44 (2.18) | 22 (2.92) |
| Stanford Hospital | California | 88 (4.35) | 45 (5.98) |
| Texas Heart Institute at Baylor | Texas | 14 (0.69) | 5 (0.66) |
| University of California, Los Angeles | California | 76 (3.76) | 29 (3.85) |
| UT Southwestern | Texas | 77 (3.81) | 38 (5.05) |
| University of Pennsylvania | Pennsylvania | 134 (6.63) | 58 (7.70) |
| Vanderbilt | Tennessee | 101 (5.00) | 29 (3.85) |

## Results

In total, 2022 cardiology physician profiles were analyzed; 37.61% (n=753) were on X, and 63.39% (n=1269) were not on X. Compared to nonusers, X users had a higher proportion of women (240/753, 31.87% vs 269/1269, 21.20%), higher academic faculty appointments, and a greater number of advanced degrees (all $P$<.001). Women and men X users had similar total practice durations (counted from fellowship training completion until 2024; median 10, IQR 1-45 y vs median 12, IQR 1-48 y; $P$=.14), but women's practice durations since joining X were significantly lower (median 6.4, IQR 5-11 y vs median 7.8, IQR 5-10 y; $P$<.001). After adjusting for the number of years on X, women and men showed similar numbers of followers (median 71.46, IQR 24.8 - 180.84 vs median 78.05, IQR 24.96 - 197.33 per year on X; $P$=.68) and posts (median 29.1, IQR 5.06 - 102.47 vs median 28.04, IQR 5.22 - 111.15 per year on X; $P$=.98), but women had higher levels of self-engagement (number of users followed: median 42.11, IQR 16.8 - 84.77 vs median 31.9, IQR 11.48 - 70.4 per year on X; $P$=.02; number of liked posts: median 112.52, IQR 16.58 - 430.1 vs median 64.49, IQR 6.94 - 318.98 per year on X; $P$=.02; Table 2). Per a thematic analysis of biographical text, women were more likely than men to mention being a parent (48/239, 20.08% vs 64/513, 12.48%; $P$=.006), but there was no significant difference in mentions of jobs ($P$=.36) or hobbies ($P$=.89; Table 2).

**Table .** Characteristics and demographics of top hospital cardiologists on X, stratified by gender.

| Variable | Not on X (n=1269) | On X (n=753) | *P* value[a] | Men on X (n=513) | Women on X (n=240) | *P* value[b] |
|---|---|---|---|---|---|---|
| Geographic region, n (%[c]) | | | <.001 | | | .72 |
| Northeast | 741 (58.39) | 364 (48.34) | | 245 (47.76) | 119 (49.58) | |
| Midwest | 249 (19.62) | 187 (24.83) | | 130 (25.34) | 57 (23.75) | |
| South | 155 (12.21) | 106 (14.08) | | 69 (13.45) | 37 (15.42) | |
| West | 124 (9.77) | 96 (12.75) | | 69 (13.45) | 27 (11.25) | |
| Gender, n (%[c]) | | | <.001 | | | —[d] |
| Men | 1000 (78.8) | 513 (68.13) | | — | — | |
| Women | 269 (21.20) | 240 (31.87) | | — | — | |
| Faculty type, n (%[c]) | | | <.001 | | | .06 |
| Not explicitly listed | 347 (27.34) | 191 (25.37) | | 135 (26.32) | 56 (23.33) | |
| Instructor/clinician | 97 (7.64) | 39 (5.18) | | 21 (4.09) | 18 (7.5) | |
| Assistant | 441 (34.75) | 227 (30.15) | | 149 (29.04) | 78 (32.5) | |
| Associate | 208 (16.39) | 153 (20.32) | | 100 (19.49) | 53 (22.08) | |
| Professor | 176 (13.87) | 143 (18.99) | | 108 (21.05) | 35 (14.58) | |
| Number of leadership titles, n (%[c]) | | | <.001 | | | .11 |
| 0 | 840 (66.19) | 360 (47.81) | | 239 (46.59) | 121 (50.42) | |
| 1 | 306 (24.11) | 241 (32.01) | | 159 (30.99) | 82 (34.17) | |
| 2 | 95 (7.49) | 111 (14.74) | | 85 (16.57) | 26 (10.83) | |
| ≥3 | 28 (2.21) | 41 (5.44) | | 30 (5.85) | 11 (4.58) | |
| Subspecialty, n (%[c]) | | | <.001 | | | <.001 |
| General | 552 (43.53) | 213 (28.29) | | 133 (25.93) | 80 (33.33) | |
| Interventional | 226 (17.82) | 112 (14.87) | | 90 (17.54) | 22 (9.17) | |
| Imaging | 193 (15.22) | 121 (16.07) | | 68 (13.26) | 53 (22.08) | |
| Congenital | 31 (2.44) | 24 (3.19) | | 12 (2.34) | 12 (5) | |
| Heart failure | 91 (7.18) | 121 (16.07) | | 78 (15.2) | 43 (17.92) | |
| Electrophysiology | 138 (10.88) | 95 (12.62) | | 84 (16.37) | 11 (4.58) | |
| Other | 37 (2.92) | 67 (8.9) | | 48 (9.36) | 19 (7.92) | |
| Dual degree, n (%[c]) | | | | | | |
| PhD | | | <.001 | | | .34 |
| No | 1183 (93.22) | 662 (87.92) | | 447 (87.13) | 215 (89.58) | |
| Yes | 86 (6.78) | 91 (12.08) | | 66 (12.87) | 25 (10.42) | |
| MS | | | <.001 | | | .71 |
| No | 1220 (96.14) | 679 (90.17) | | 464 (90.45) | 215 (89.58) | |
| Yes | 49 (3.86) | 74 (9.83) | | 49 (9.55) | 25 (10.42) | |
| MPH | | | <.001 | | | .55 |
| No | 1226 (96.61) | 680 (90.31) | | 461 (89.86) | 219 (91.25) | |
| Yes | 43 (3.39) | 73 (9.69) | | 52 (10.14) | 21 (8.75) | |
| MBA | | | .24 | | | .76 |
| No | 1255 (98.9) | 740 (98.27) | | 503 (98.05) | 237 (98.75) | |
| Yes | 14 (1.1) | 13 (1.73) | | 10 (1.95) | 3 (1.25) | |

| Variable | Not on X (n=1269) | On X (n=753) | *P* value[a] | Men on X (n=513) | Women on X (n=240) | *P* value[b] |
|---|---|---|---|---|---|---|
| Practice duration (years) | | | <.001 | | | .14 |
| Median (IQR) | 21 (12 - 31) | 11 (6 - 21) | | 12 (1 - 48) | 10 (1 - 45) | |
| Overall: <9; physicians on X: <7, n (%[e]) | 152 (37.91) | 249 (62.09) | | 117 (68.82) | 53 (31.18) | |
| Overall: ≥9 and <17; physicians on X: ≥7 and <11, n (%[e]) | 270 (60.81) | 174 (39.19) | | 153 (63.22) | 89 (36.78) | |
| Overall: ≥17 and <28; physicians on X: ≥11 and <21, n (%[e]) | 292 (67.13) | 143 (32.87) | | 123 (69.49) | 54 (30.51) | |
| Overall: ≥28; physicians on X: ≥21, n (%[e]) | 359 (80.86) | 85 (19.14) | | 120 (73.17) | 44 (26.83) | |
| X use variables (publicly available), median (IQR) | | | | | | |
| Time on X (years) | — | — | — | 7.80 (5.30 - 11.34) | 6.39 (5.06 - 10.11) | <.001 |
| Average number of followers per year on X | — | — | — | 78.05 (24.96 - 197.33) | 71.46 (24.8 - 180.84) | .68 |
| Average number of people followed per year on X | — | — | — | 31.90 (11.48 - 70.40) | 42.11 (16.8 - 84.77) | .02 |
| Average number of tweets per year on X | — | — | — | 28.04 (5.22 - 111.15) | 29.10 (5.06 - 102.47) | .98 |
| Average number of media posts per year on X | — | — | — | 2.27 (0.26 - 10.38) | 2.20 (0.26 - 10.78) | .96 |
| Average number of liked posts per year on X | — | — | — | 64.49 (6.94 - 318.98) | 112.52 (16.58 - 430.1) | .02 |
| Thematic content of X biography, n (%[c]) | | | | | | |
| Job Roles | | | — | | | .36 |
| No mention | — | — | | 98 (19.10) | 39 (16.32) | |
| Mention | — | — | | 415 (80.90) | 200 (83.68) | |
| Specialty | | | — | | | .48 |
| No mention | — | — | | 169 (32.94) | 85 (35.56) | |
| Mention | — | — | | 344 (67.06) | 154 (64.44) | |
| Parent | | | — | | | .006 |
| No mention | — | — | | 449 (87.52) | 191 (79.92) | |
| Mention | — | — | | 64 (12.48) | 48 (20.08) | |
| Spouse | | | — | | | .77 |
| No mention | — | — | | 467 (91.03) | 216 (90.38) | |
| Mention | — | — | | 46 (8.97) | 23 (9.62) | |
| Institution | | | — | | | .56 |
| No mention | — | — | | 148 (28.85) | 64 (26.78) | |

| Variable | Not on X (n=1269) | On X (n=753) | *P* value[a] | Men on X (n=513) | Women on X (n=240) | *P* value[b] |
|---|---|---|---|---|---|---|
| Mention | — | — | | 365 (71.15) | 175 (73.22) | |
| Personal interests | | | — | | | .89 |
| No mention | — | — | | 444 (86.55) | 206 (86.19) | |
| Mention | — | — | | 69 (13.45) | 33 (13.81) | |

[a]Not on X versus on X. The *P* values were calculated via Wilcoxon rank-sum tests for continuous and ordinal variables and via chi-square tests or Fisher exact tests for categorical variables, as appropriate.

[b]Men on X versus women on X. The *P* values were calculated via Wilcoxon rank-sum tests for continuous and ordinal variables and via chi-square tests or Fisher exact tests for categorical variables, as appropriate.

[c]Column %: these percentages were calculated based on the total n values for the columns of this section.

[d]Not applicable.

[e]Row %: these percentages were calculated based on the total n values for the rows of this section.

## Discussion

In our analysis of U.S. News Best Hospitals cardiologists, the proportion of women on X was higher than the proportion of women non–X users. One possible explanation for this is that women cardiologists may be seeking novel opportunities for networking, collaboration, visibility, and/or self-promotion that are not available through traditional channels [5]. Additionally, compared to men, women cardiologists had similar time-adjusted follower counts but liked more posts. This is consistent with content language analyses demonstrating higher expected levels of friendliness in women's professional communications, including more frequent use of exclamation points as markers of friendly interaction, which is associated with increased emotional labor [8,9]. Further, women cardiologists were more likely to mention being a parent, suggesting that women may be more comfortable with highlighting work-life integration factors. This is unsurprising, as women physicians have joined social media groups discussing issues such as parenting, maternity leave, and women leadership in medicine [5]. These observations support efforts to better understand motivational differences in social media use and impacts on potential downstream professional benefits.

Our study has several limitations, including institutional websites being subject to inaccuracy and incompleteness, currently available X data being more limited compared to prior studies, limited physician practice type information, and potential misgendering [10]. However, our findings highlight the increased presence of women cardiologists on X, with similar influence to men and higher engagement despite shorter time on X. These findings suggest an inherent desire to engage on social media for professional use, though the motivating factors driving these behavioral differences and their impact on existing gender disparities warrant further study.

### Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

### Authors' Contributions

Conceptualization: KMA, APN, MS
Data curation: MS, SK, HT, OP
Formal analysis: MS, SK
Investigation: MS, SK, HT, OP
Methodology: KMA, SK, MS, MK
Project administration: KMA, APN
Resources: KMA, SK
Software: SK
Supervision: KMA, APN
Validation: KMA, SK
Visualization: KMA, APN, MK, SK, MS
Writing – original draft: KMA, MS
Writing – review & editing: KMA, APN, MS, MK

### Conflicts of Interest

KMA reports honoraria from OncLive, outside of the submitted work. The remaining authors have no disclosures.

### References

1. Murphy M, Callander JK, Dohan D, Grandis JR. Networking practices and gender inequities in academic medicine: women's and men's perspectives. EClinicalMedicine 2022 Mar 13;45:101338. [doi: 10.1016/j.eclinm.2022.101338] [Medline: 35299655]
2. Yong CM, Balasubramanian S, Douglas PS, et al. Temporal trends in the proportion of women physician speakers at major cardiovascular conferences. Circulation 2021 Feb 16;143(7):755-757. [doi: 10.1161/CIRCULATIONAHA.120.052663] [Medline: 33587663]
3. Shahid I, Khan MS, Sohail A, et al. Evaluation of representation of women as authors in pivotal trials supporting US Food and Drug Administration approval of novel cardiovascular drugs. JAMA Netw Open 2022 Feb 1;5(2):e220035. [doi: 10.1001/jamanetworkopen.2022.0035] [Medline: 35212753]
4. Denby KJ, Szpakowski N, Silver J, Walsh MN, Nissen S, Cho L. Representation of women in cardiovascular clinical trial leadership. JAMA Intern Med 2020 Oct 1;180(10):1382-1383. [doi: 10.1001/jamainternmed.2020.2485] [Medline: 32865557]
5. Shillcutt SK, Silver JK. Social media and advancement of women physicians. N Engl J Med 2018 Jun 14;378(24):2342-2345. [doi: 10.1056/NEJMms1801980] [Medline: 29897857]
6. Chandra NV, Watson KE, Han JK. If you can't see it, you can't be it: mentorship for female internal medicine residents. JACC Case Rep 2020 May 20;2(5):840-842. [doi: 10.1016/j.jaccas.2020.03.009] [Medline: 34317362]
7. Best hospitals for cardiology, heart & vascular surgery. US News & World Report. 2024. URL: https://health.usnews.com/best-hospitals/rankings/cardiology-and-heart-surgery [accessed 2025-05-09]
8. Waseleski C. Gender and the use of exclamation points in computer-mediated communication: an analysis of exclamations posted to two electronic discussion lists. J Comput Mediat Commun 2006 Jul 1;11(4):1012-1024. [doi: 10.1111/j.1083-6101.2006.00305.x]
9. Vial AC, Cowgill CM. Heavier lies her crown: gendered patterns of leader emotional labor and their downstream effects. Front Psychol 2022 Aug 29;13:849566. [doi: 10.3389/fpsyg.2022.849566] [Medline: 36106035]
10. Zhu JM, Pelullo AP, Hassan S, Siderowf L, Merchant RM, Werner RM. Gender differences in Twitter use and influence among health policy and health services researchers. JAMA Intern Med 2019 Dec 1;179(12):1726-1729. [doi: 10.1001/jamainternmed.2019.4027] [Medline: 31609393]



*Edited by A Coristine; submitted 09.09.24; peer-reviewed by A Volgman, I Kedan; revised version received 02.04.25; accepted 14.04.25; published 04.06.25.*

*Please cite as:*
*Seok M, Kim S, Tzou H, Peony O, Kamrava M, Nikolova AP, Atkins KM*
*Gender Differences in X (Formerly Twitter) Use, Influence, and Engagement Among Cardiologists From the Top U.S. News Best Hospitals*
*JMIR Cardio 2025;9:e66308*
*URL: https://cardio.jmir.org/2025/1/e66308*
*doi:10.2196/66308*

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org