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

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.

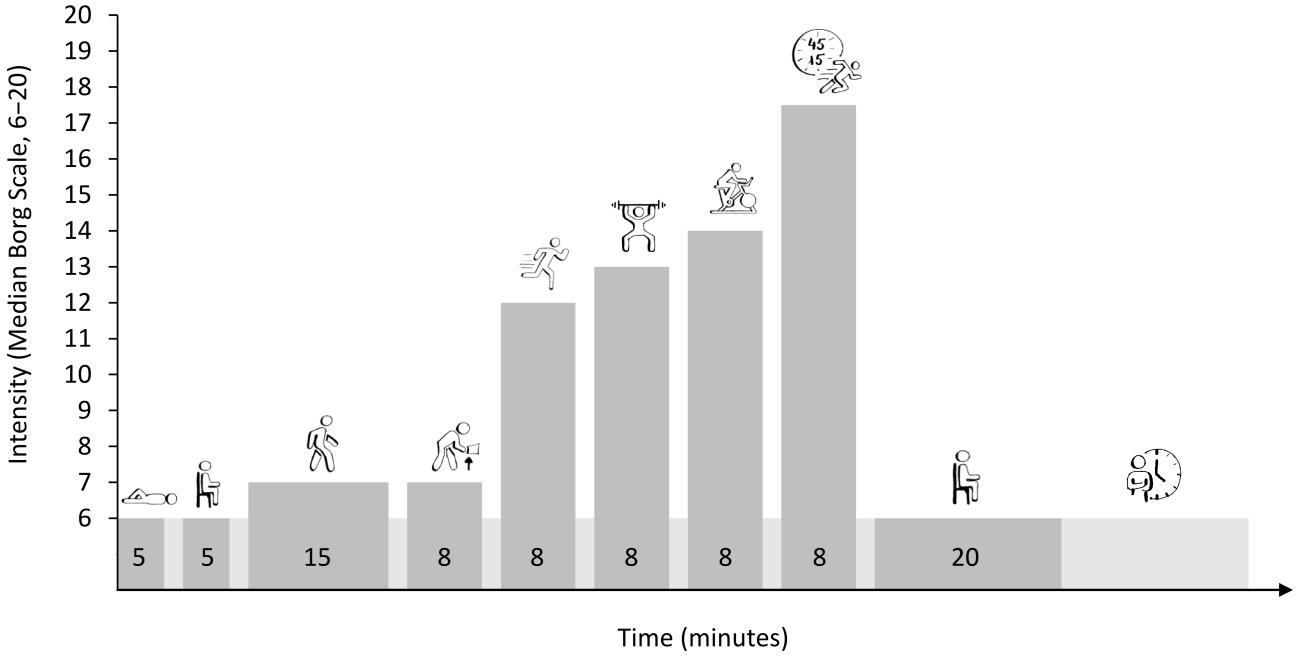
Experimental Procedure

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

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

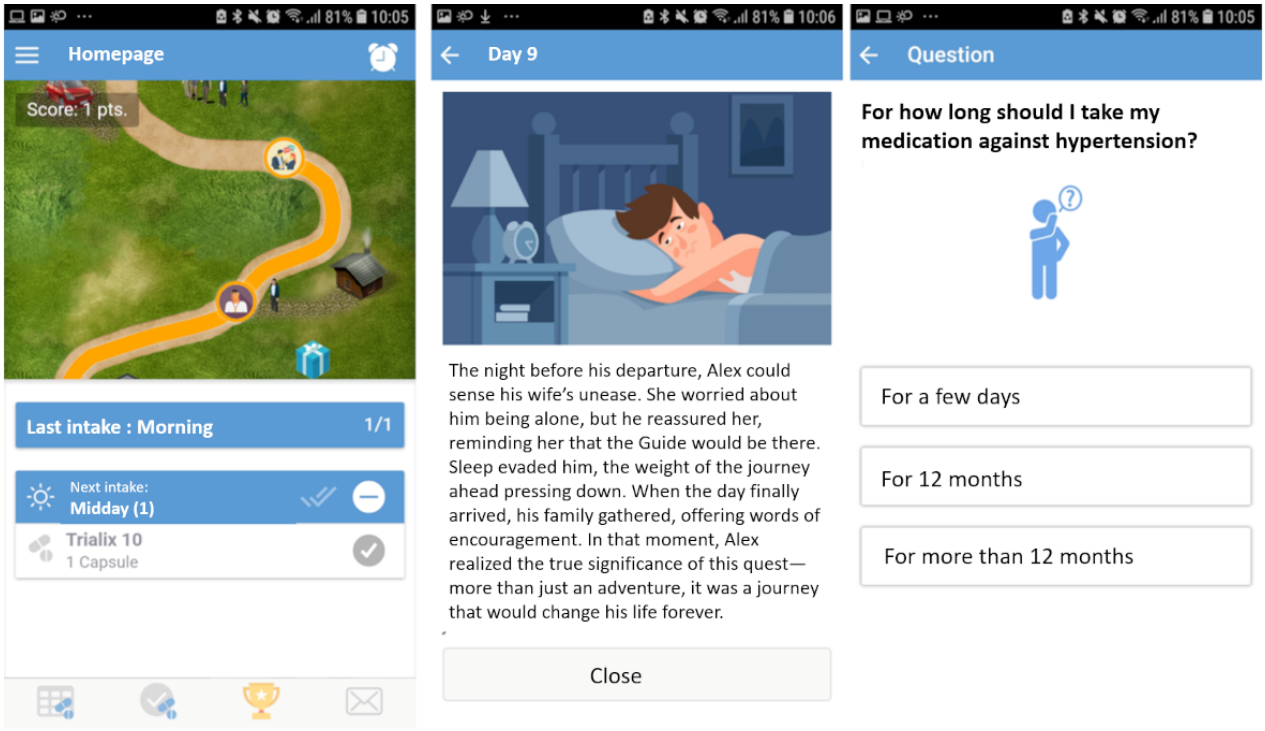
Smart-Meds is an app created following a participatory design. Users were involved all along its development, providing feedback at each step of the iterative cycles of formative evaluation [18]. The users participating in the app conception were patients participating in or having recently completed 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.

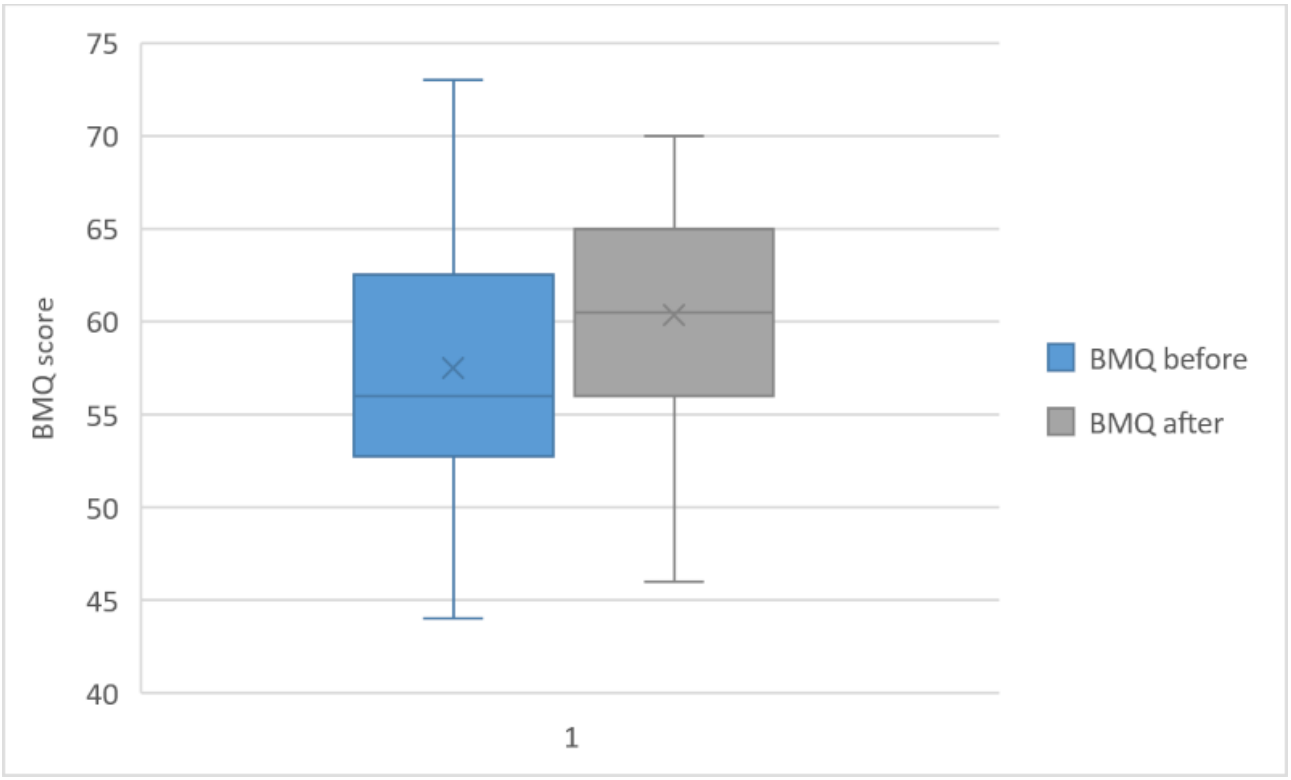
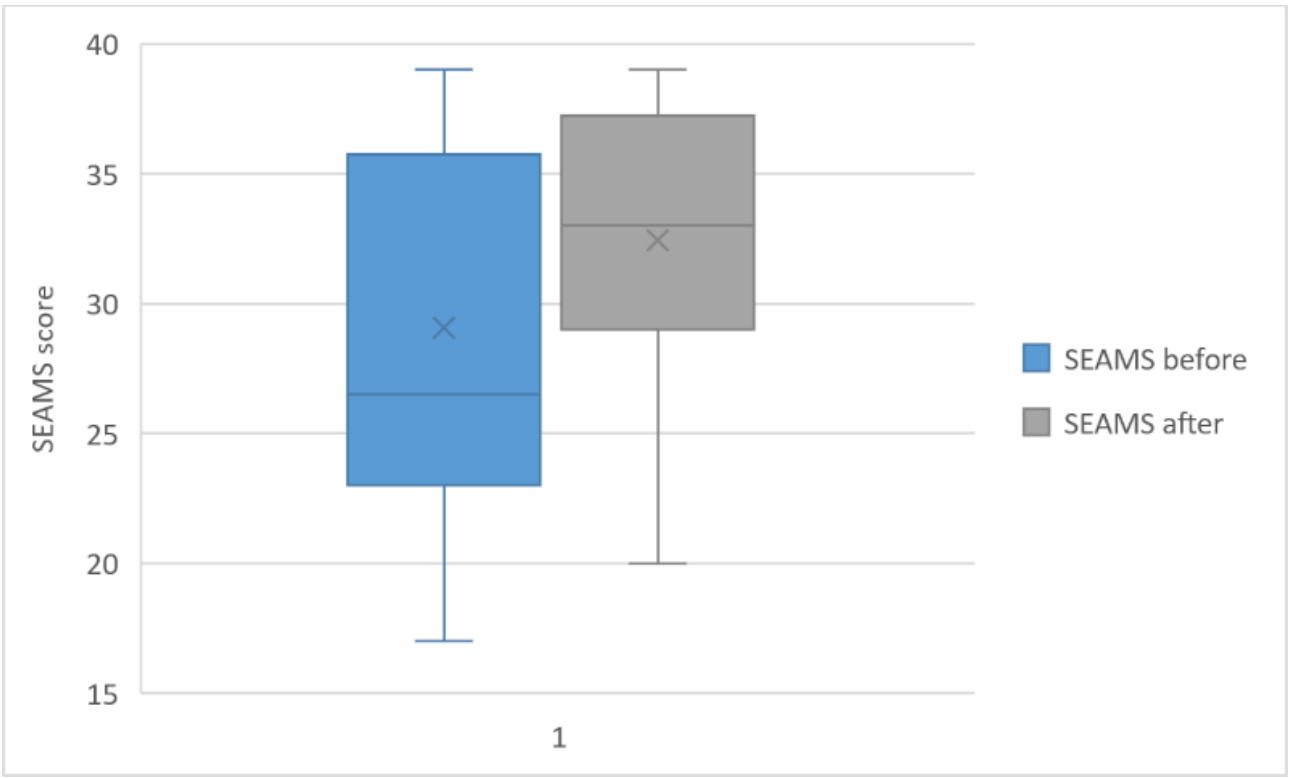


Figure 3. Boxplot of the SEAMS score before and after the intervention period for the 18 participants. SEAMS: Self-Efficacy for Appropriate Medication Use Scale.



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.

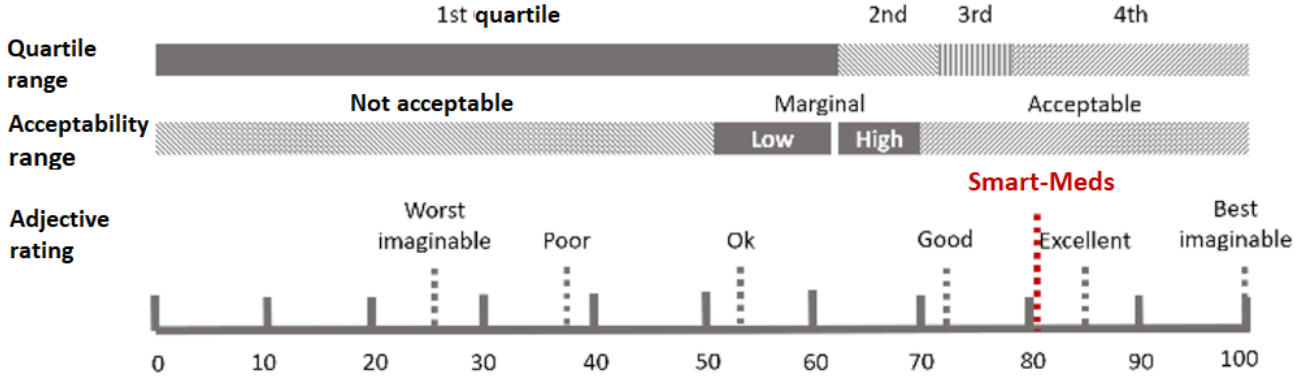
Of the 18 participants, 17 (94%) tested the quizzes and 15 (83%) enjoyed challenging their knowledge about their disease and their medications in this manner. In fact, 1 participant even suggested adding a reminder to take the quiz. Opinions about the motivational story were more varied because many 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

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.

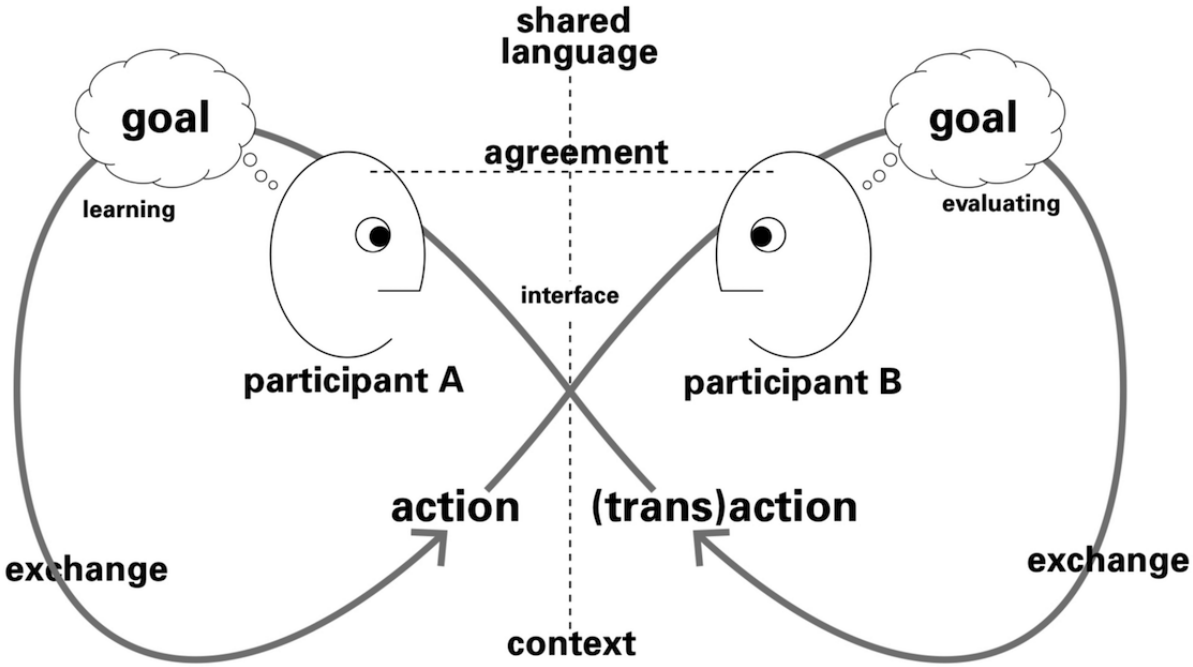
A CA is a computer program capable of understanding natural human language (in text, speech, or both forms) and responding autonomously using the same language [10]. They can be accessed through a variety of ways, such as social media platforms (eg, Facebook Messenger), websites, and smartphone apps, 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.

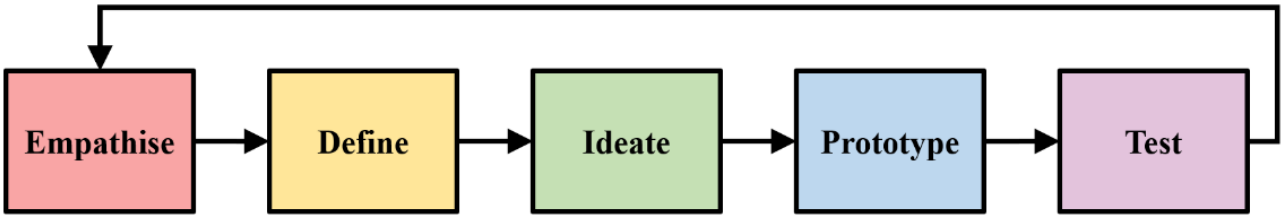
Methods

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

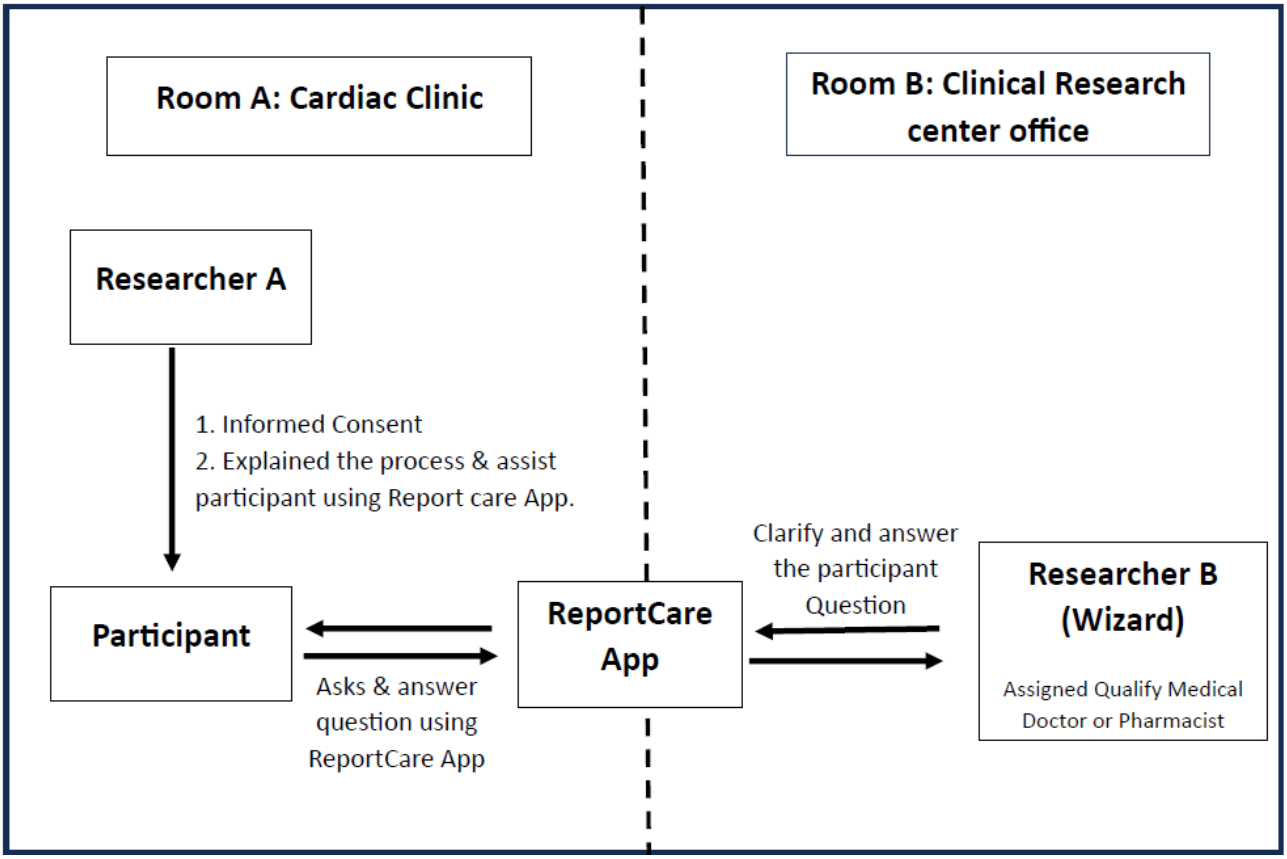
Participant names for this research have been deidentified and linked only with a study identification number. Therefore, the research did not identify the participant's identity and instead used anonymized identification numbers on all the data sets. All data are stored in Monash University Malaysia 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

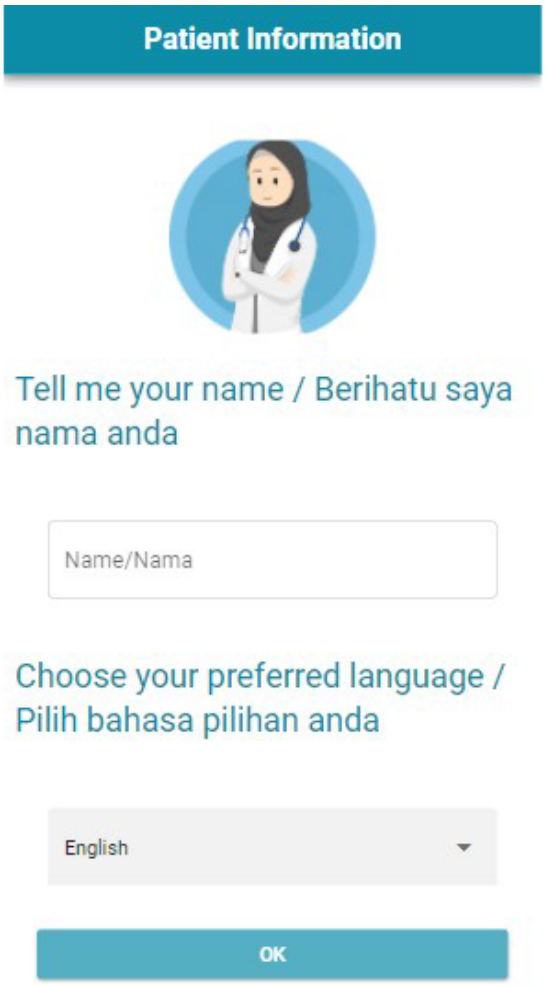
Overview

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).

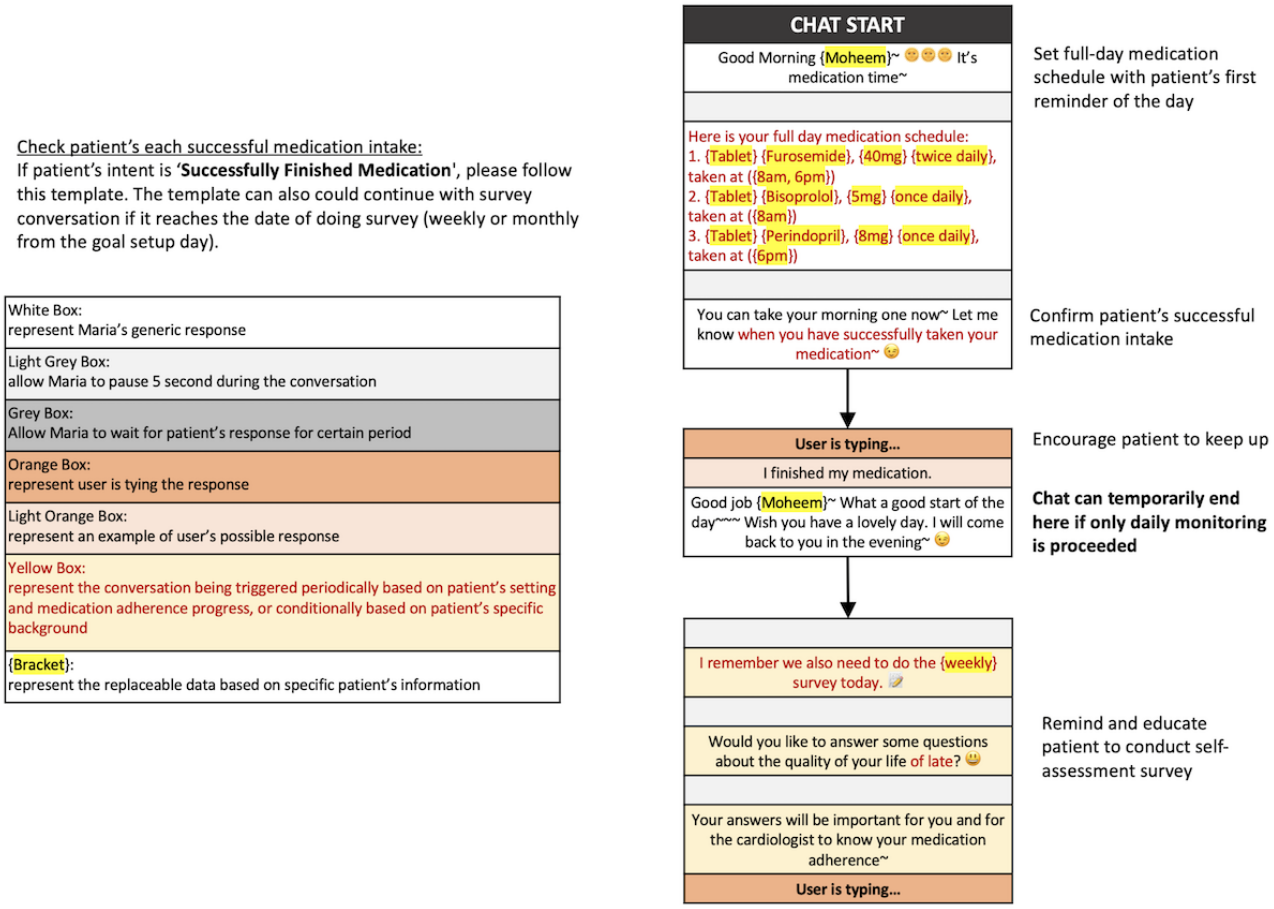
Figure 4. Screenshot of the app displaying the log-in interface including the language selection feature.



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

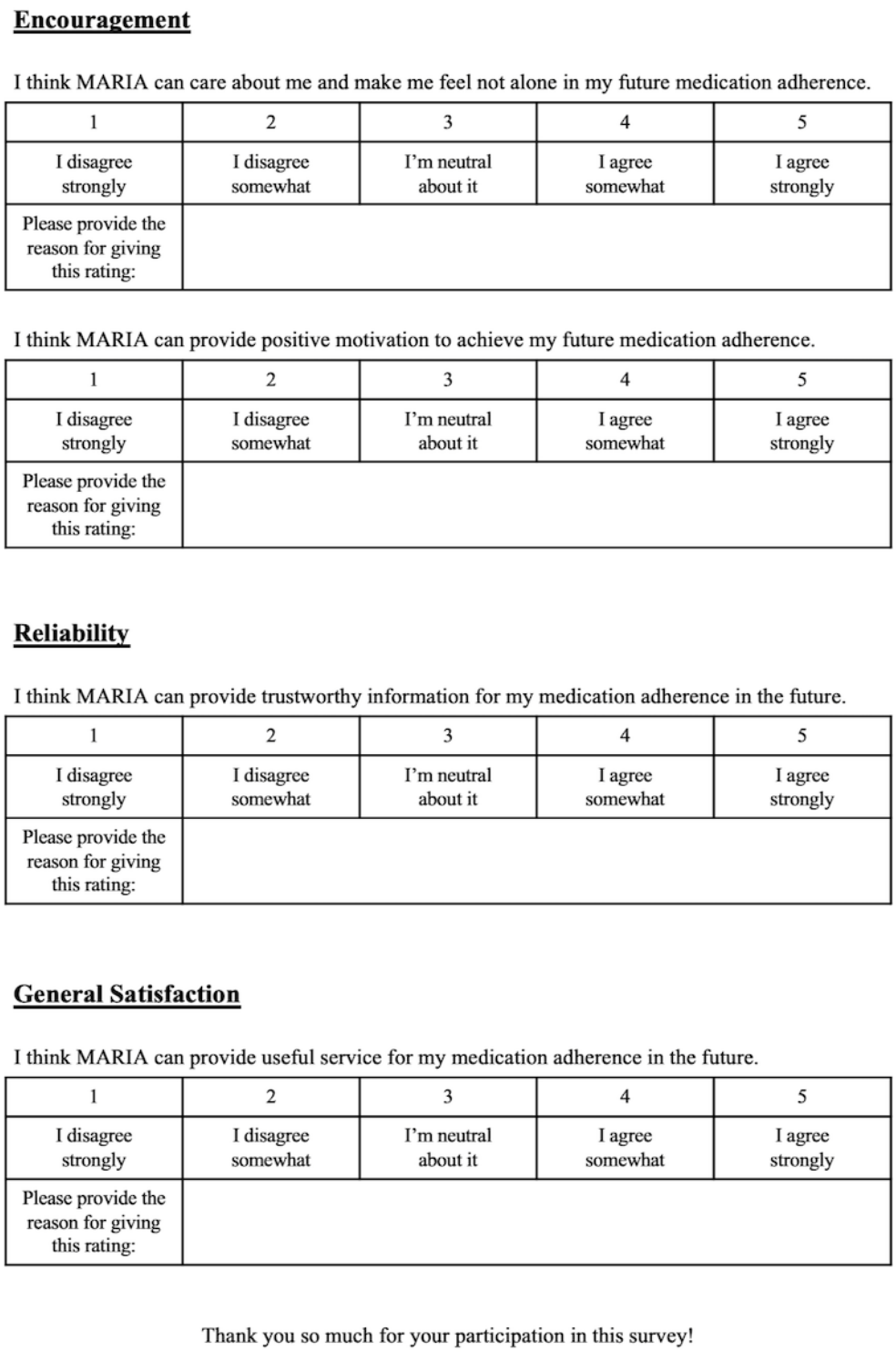
Speaker 1: "I'm having some people to dinner on Saturday, and I'd really like you to come."

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:

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].

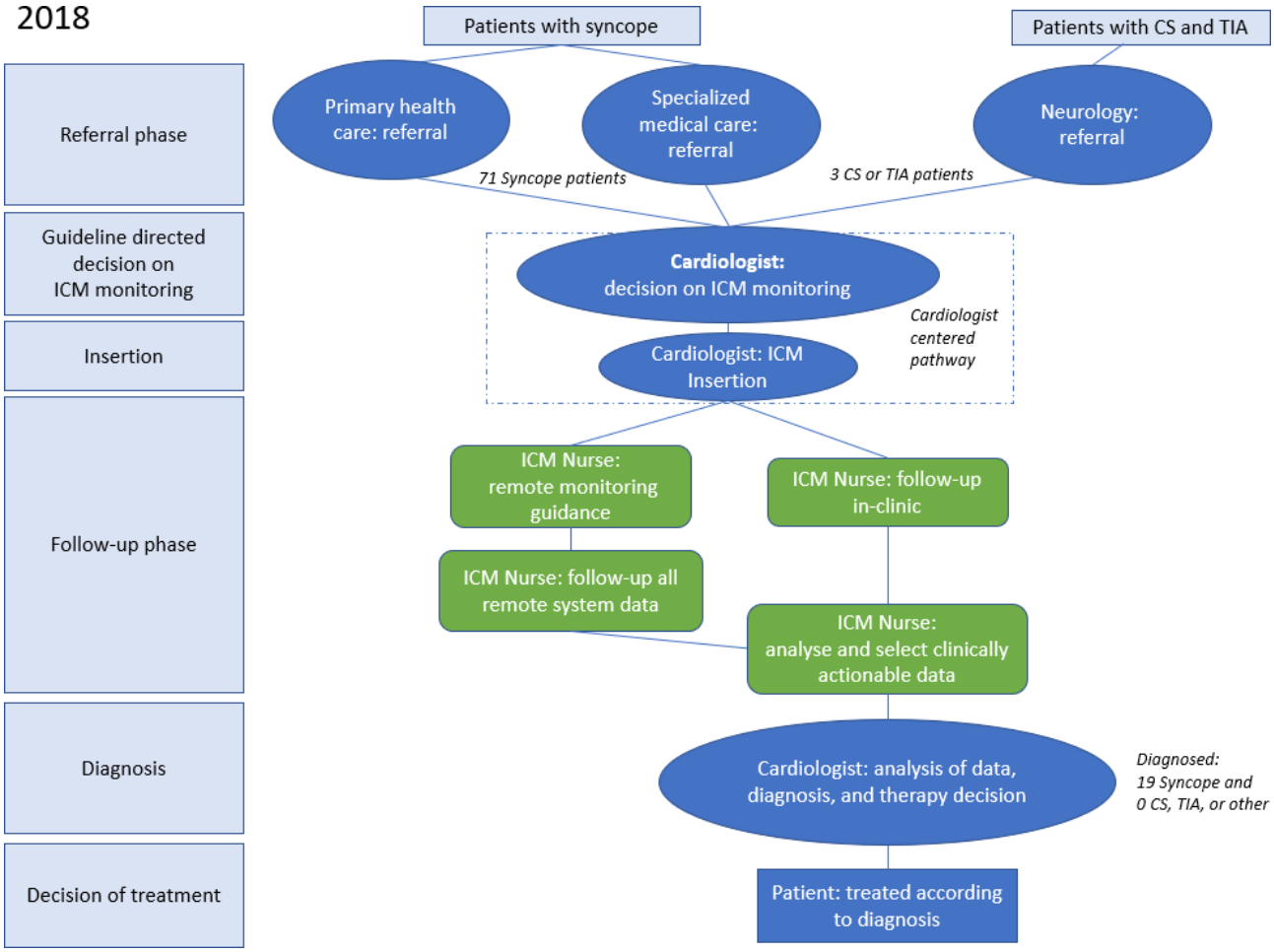
At the Finnish Tampere Heart Hospital, both insufficient staff resources and a growing number of patients in need of ICM monitoring led to the restructuring of the clinical patient pathway. The changes centered around training nurses to perform ICM insertions, the inclusion of the neurology department 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.

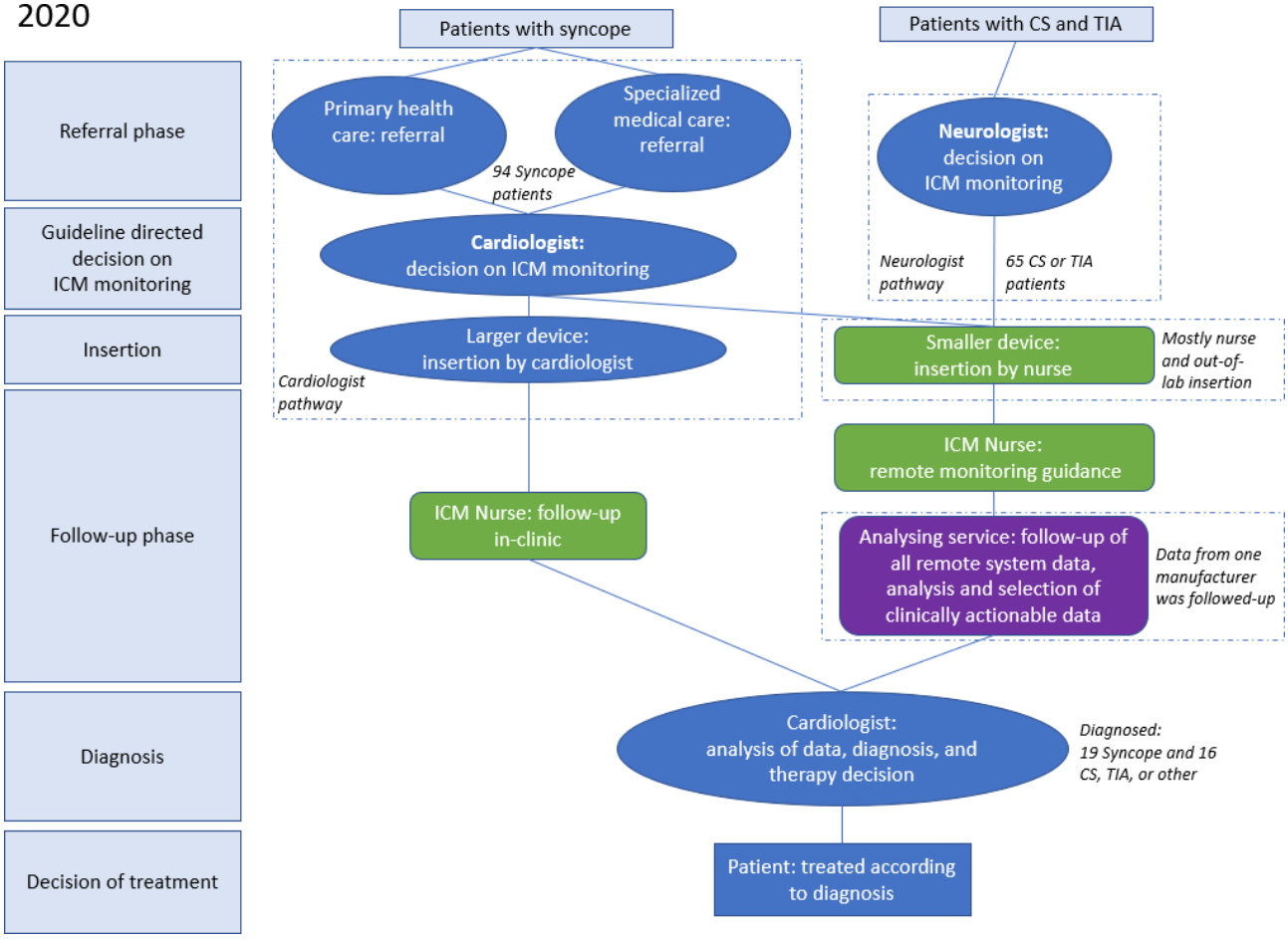


Changes in the ICM Pathway as of 2020

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.

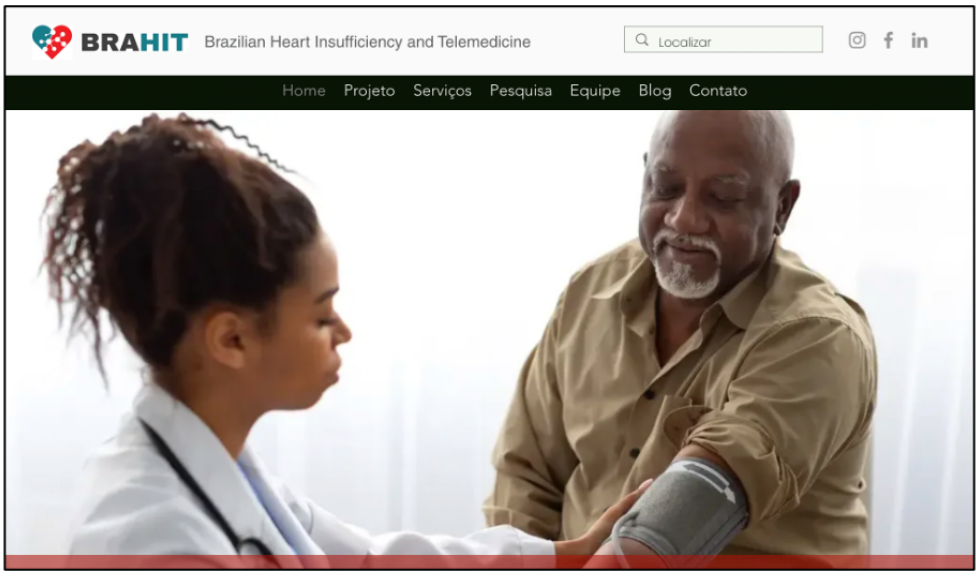
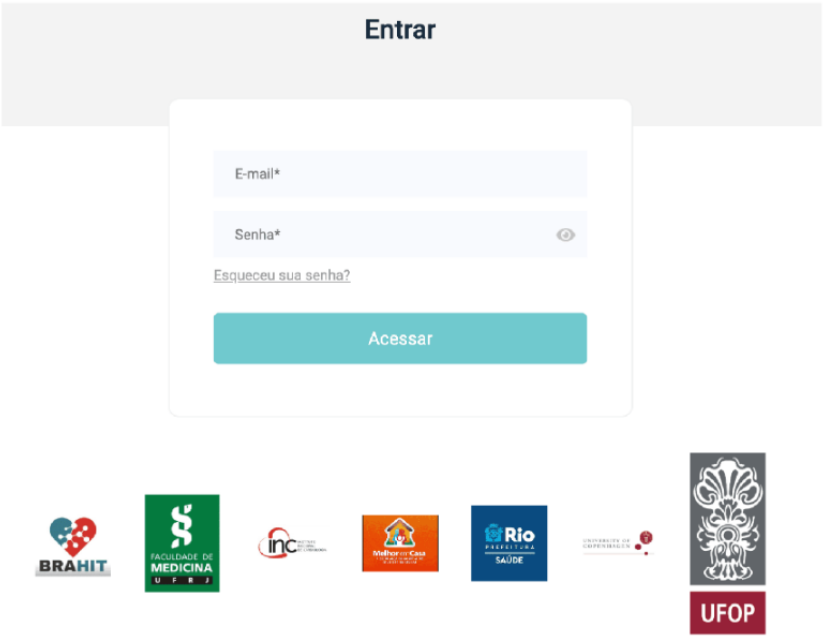


Figure 2. Log-in page for the online platform, restricted to registered users to protect data access and ensure their safety.



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.

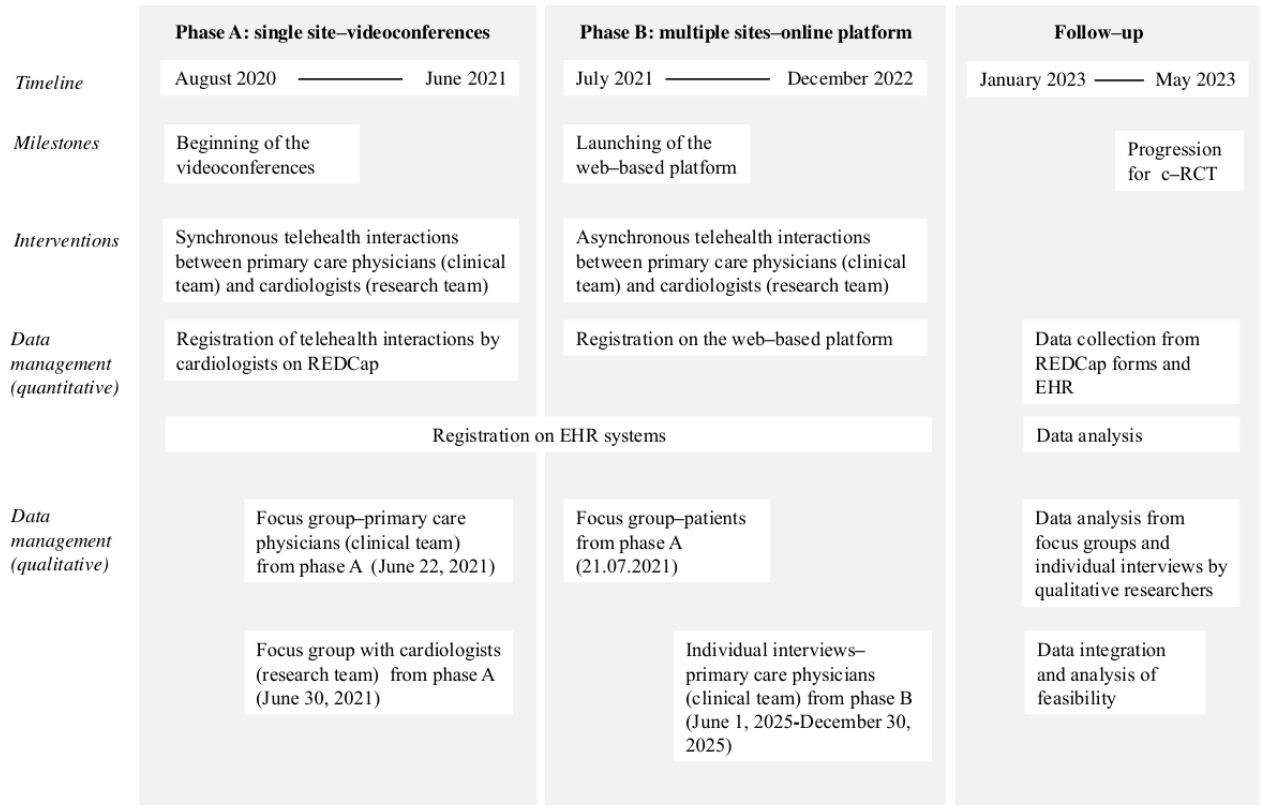
Registration

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]

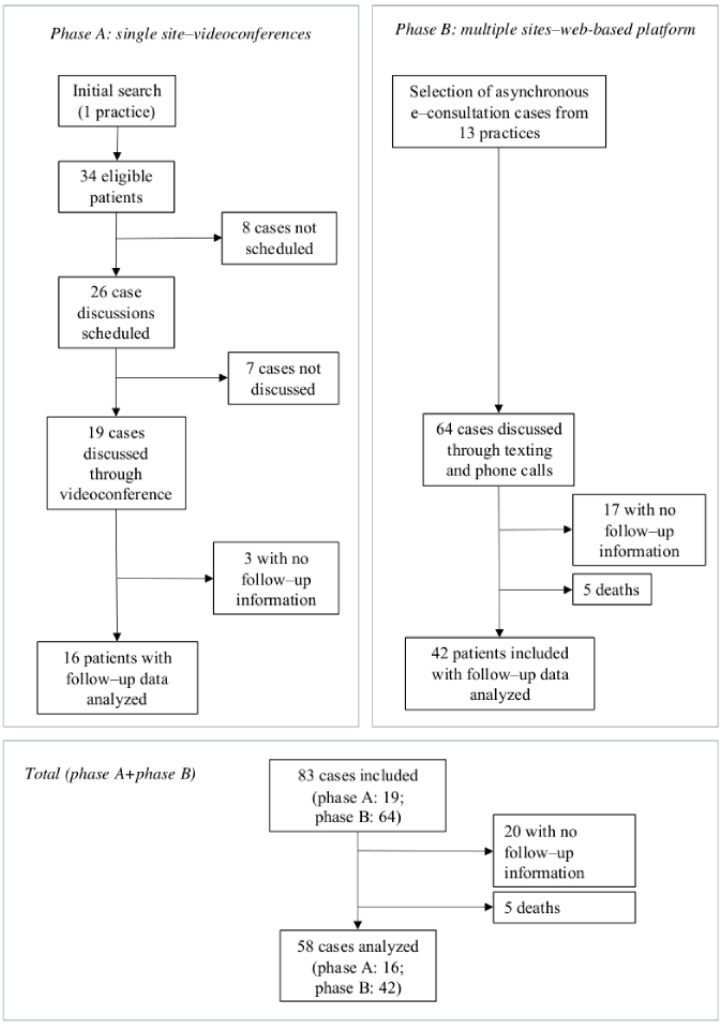
...The patients have already been managed via referral through the referral system. [IP8]

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. • 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. • 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. • Facing difficulties, physicians may privilege patients with easier access to care. • Actions integrated with telehealth support aimed at patient health literacy could be synergic.	• Adapt the intervention for the setting conditions. • Be aware of possible access hardships for non-White populations. • 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. • 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. • Work overload hindered the use of cardiologist support with telehealth.	• The results agree and are likely to have a strong correlation. • 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. • 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 • Use rate of newer agents to treat heart failure improved from 0 (0%) to 5 (20%). • 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. • 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. • Design cointerventions to overcome barriers (eg, professional education activities).
Intervention delivery	• Identification of improvements related to the intervention • 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 • 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. • The telehealth offer seemed cost-effective and did not cause a burden to the project finances.	• The feedback from teleconsultants was positive. • 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:

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.

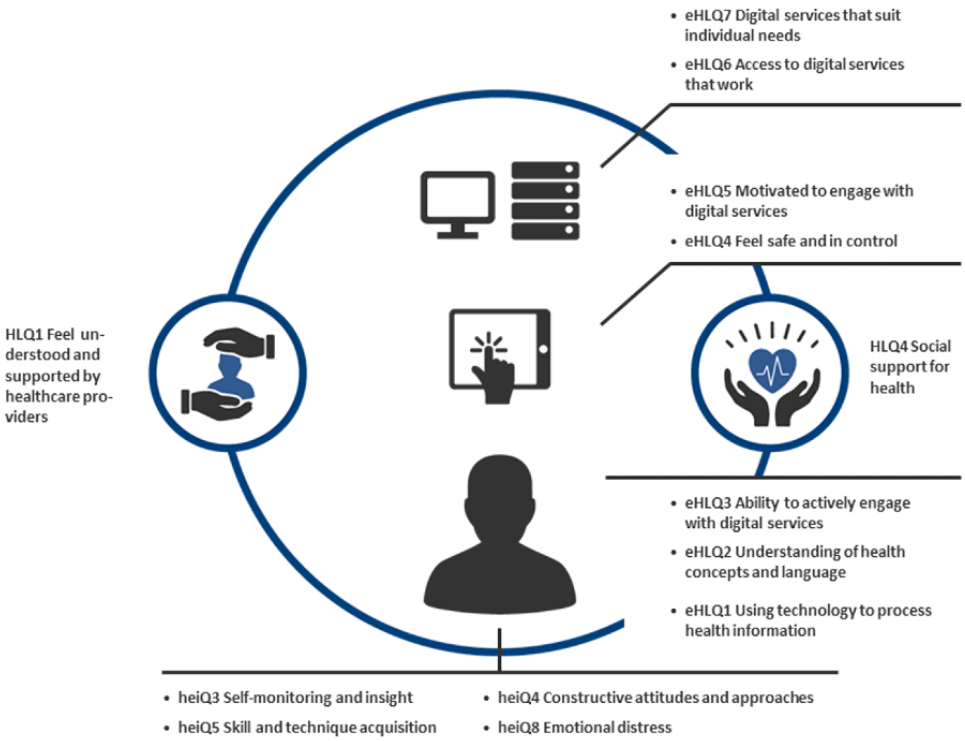
Setting, Recruitment, and Participants

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 (η^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, η^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, η^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

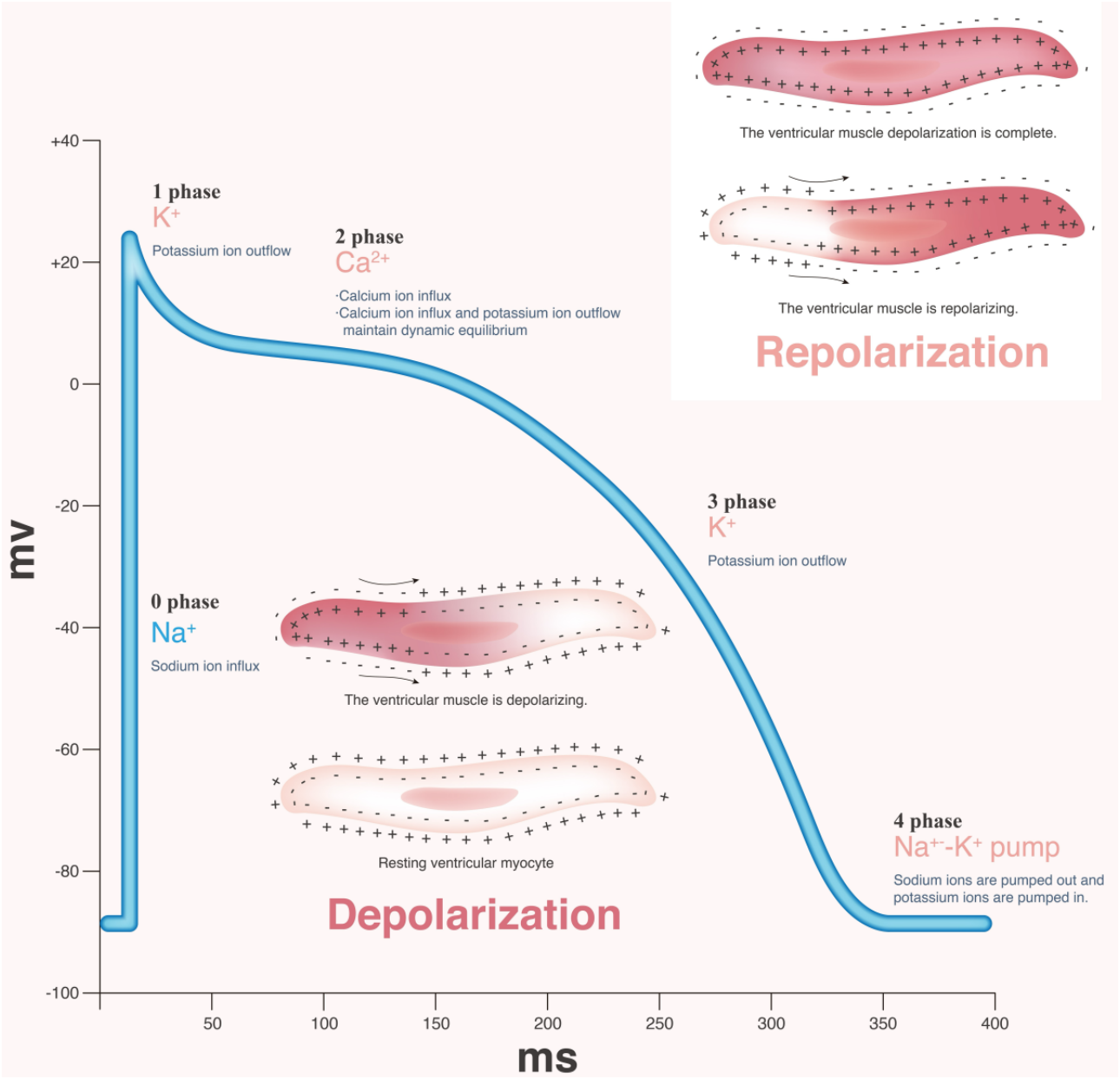
the frog nervous system. More researchers followed him, and in 1902, Einthoven broke new ground by accurately recording the electrical activity of the heart using his string galvanometer [8,9]. The string galvanometer was not without its drawbacks; it required the patient to place their hands and 1 foot 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

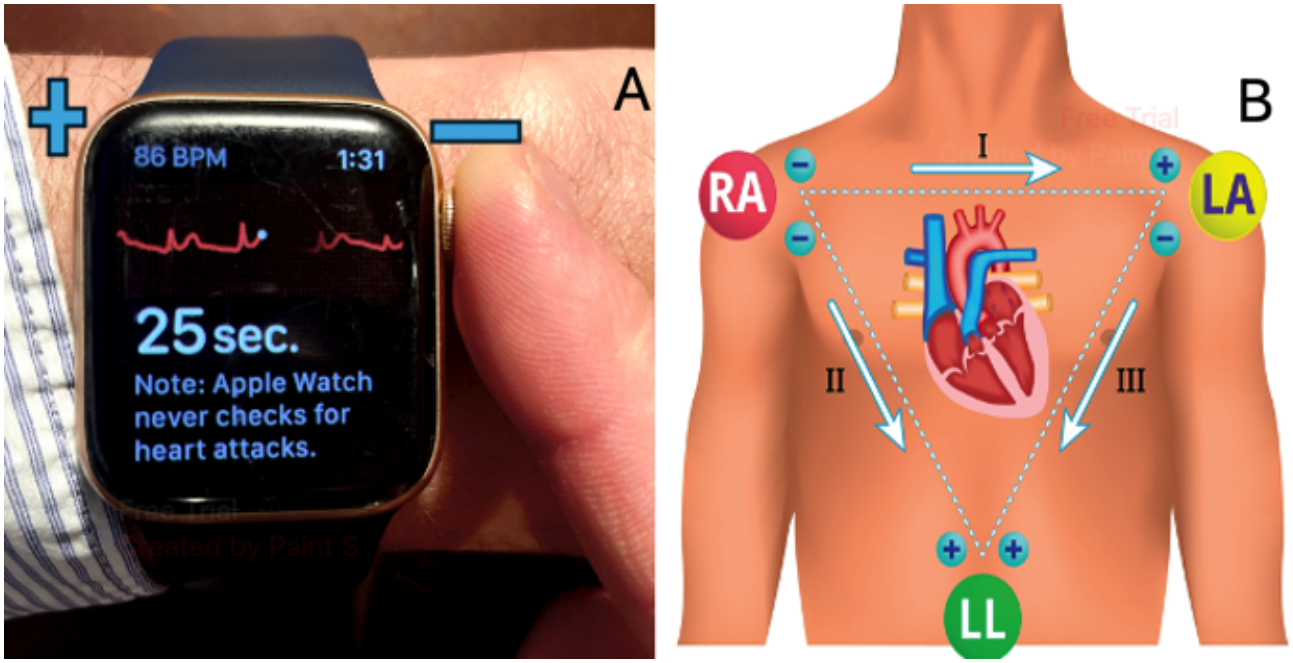
placement of 10 separate electrodes to create an electrical window for each lead [2].

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.

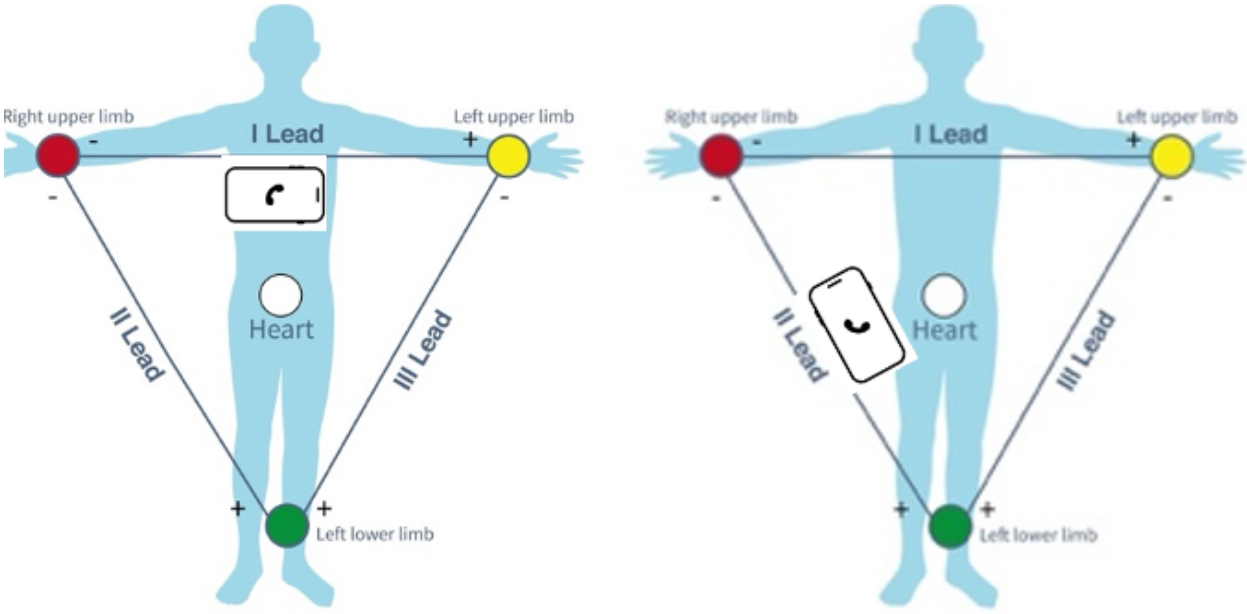
The reasons for these potential benefits over existing methodologies of AF screening and diagnosis have already been discussed; some of the biggest advantages are that patients are more likely to wear these comfortable, easily accessible devices, faster ECG analysis using AI algorithms with increasing 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

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.

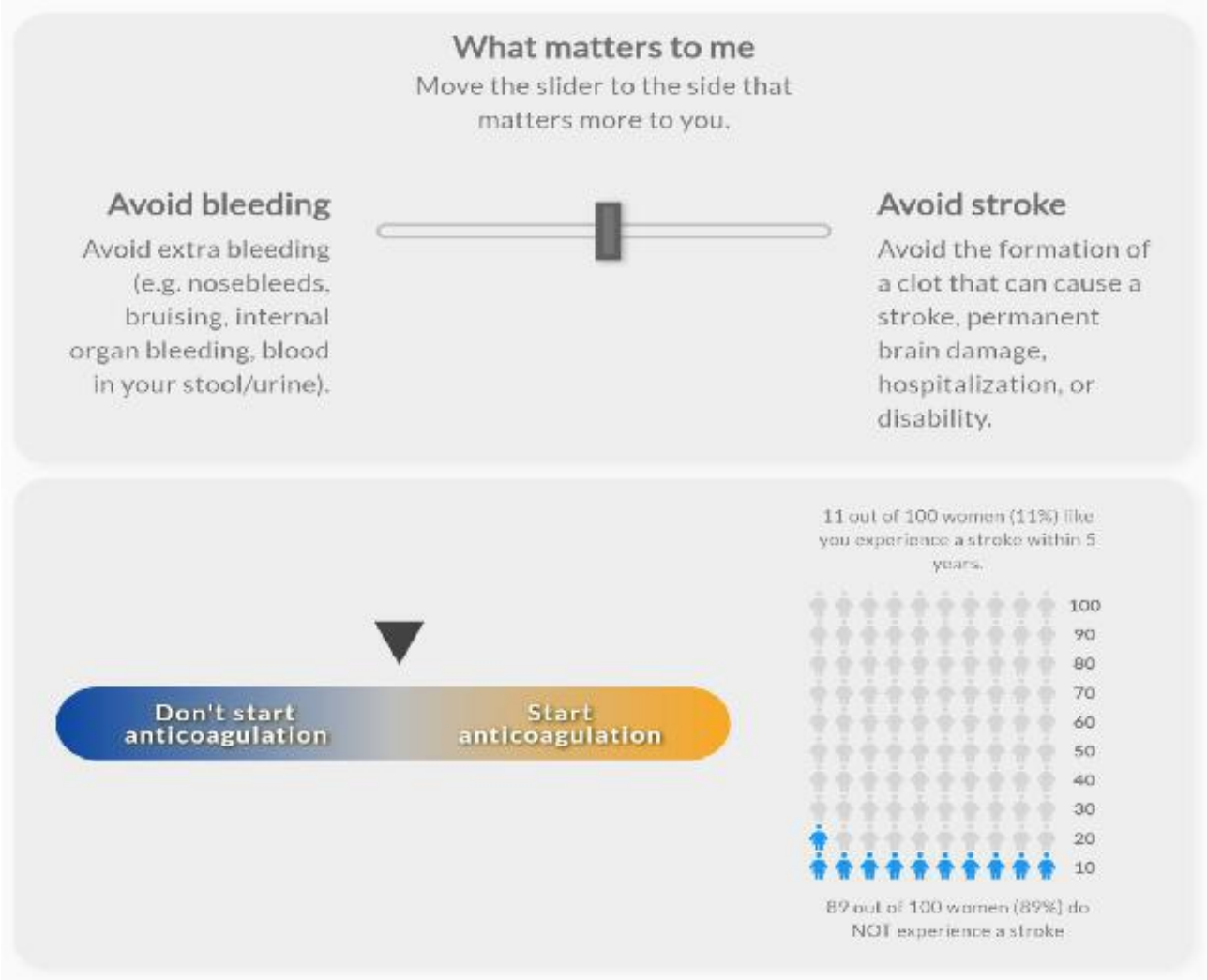


Figure 2. Example visualization of values clarification with gauge for a 75-year-old female with hypertension.

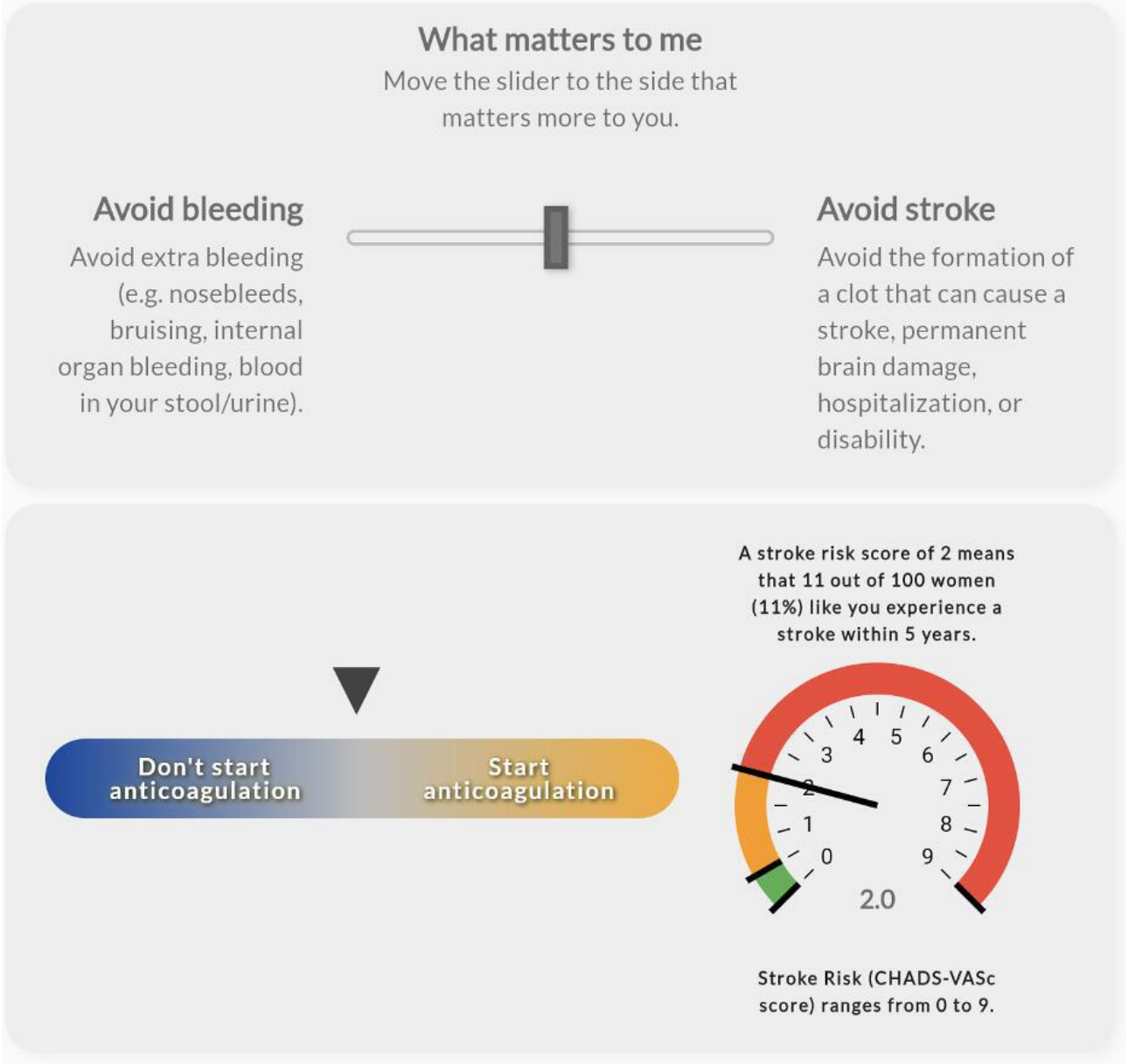


Figure 3. Example visualization with icon array for a 75-year-old female with hypertension.

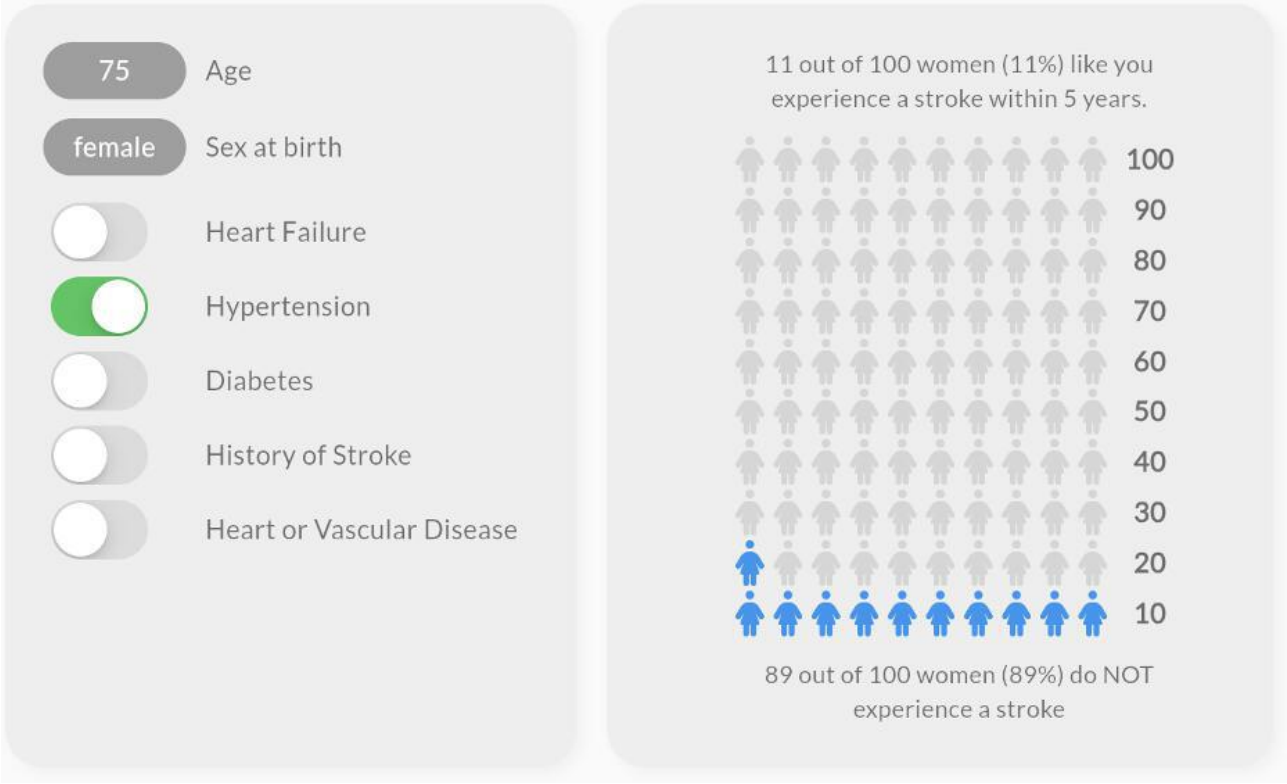


Figure 4. Example visualization with gauge for a 75-year-old female with hypertension.

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.

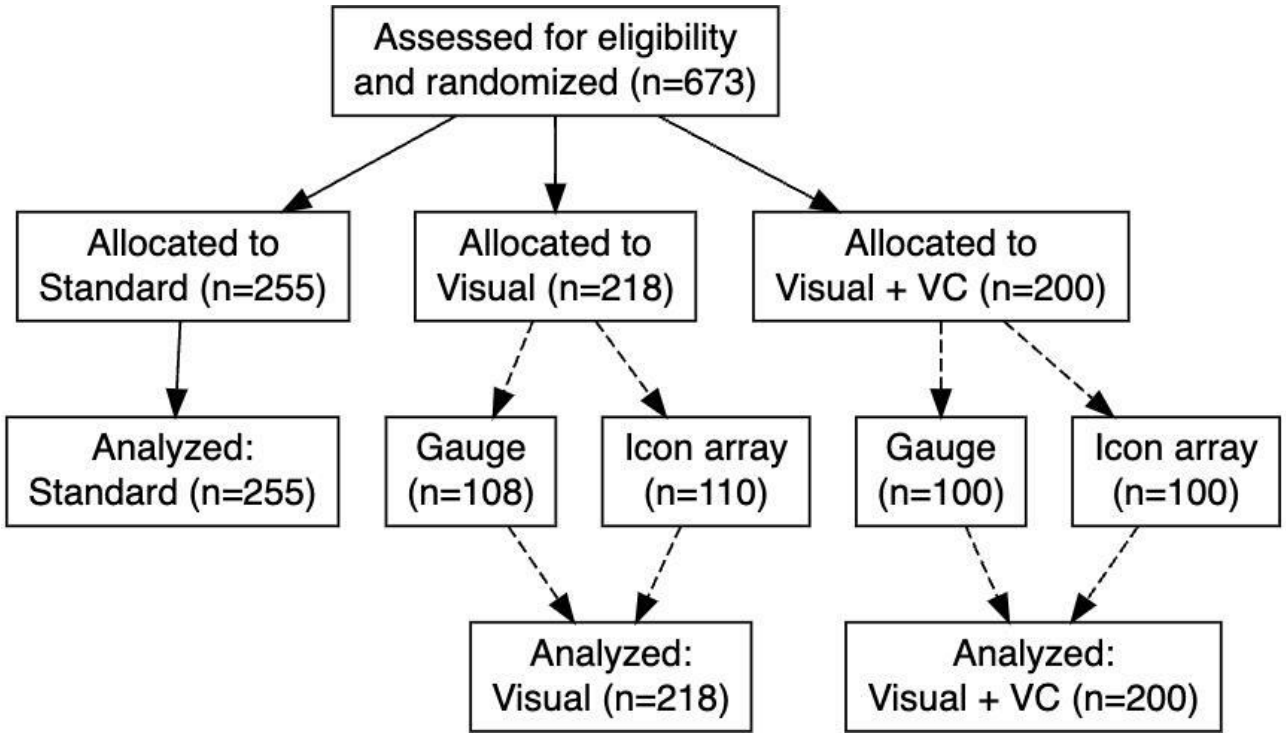


Table . Baseline characteristics.

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 • 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 • 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 • 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 • 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 • 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.

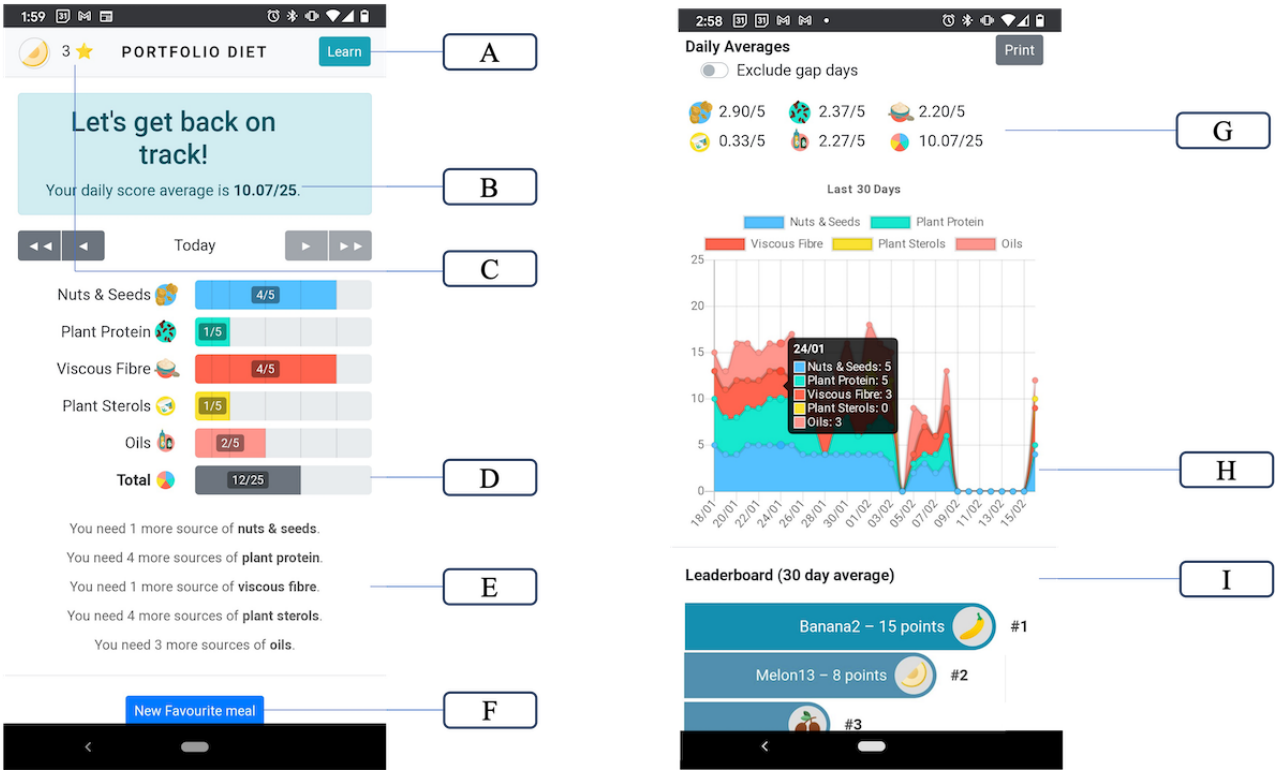
Theoretical and Operational Design Components of the PortfolioDiet.app

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.

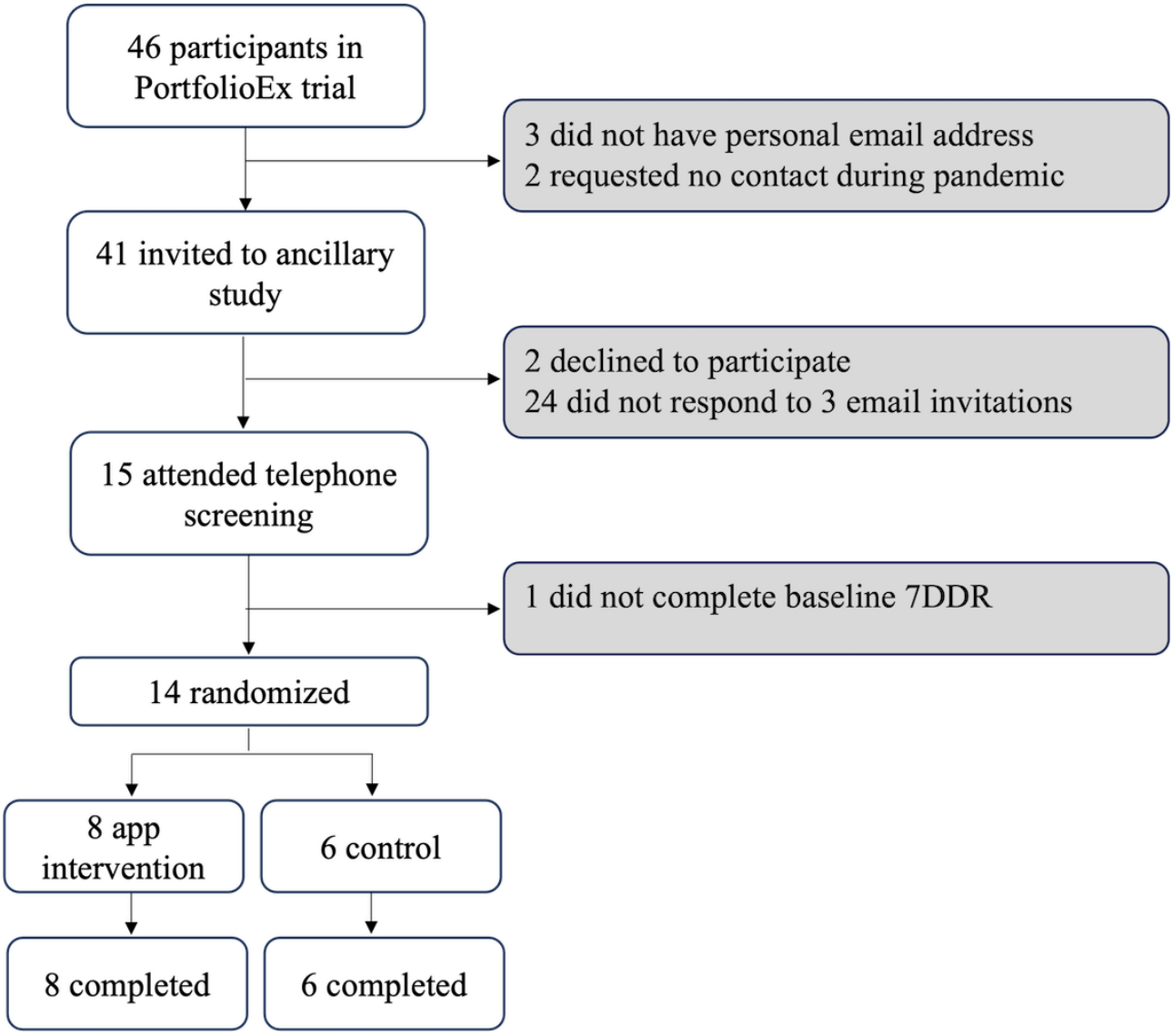
Results

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

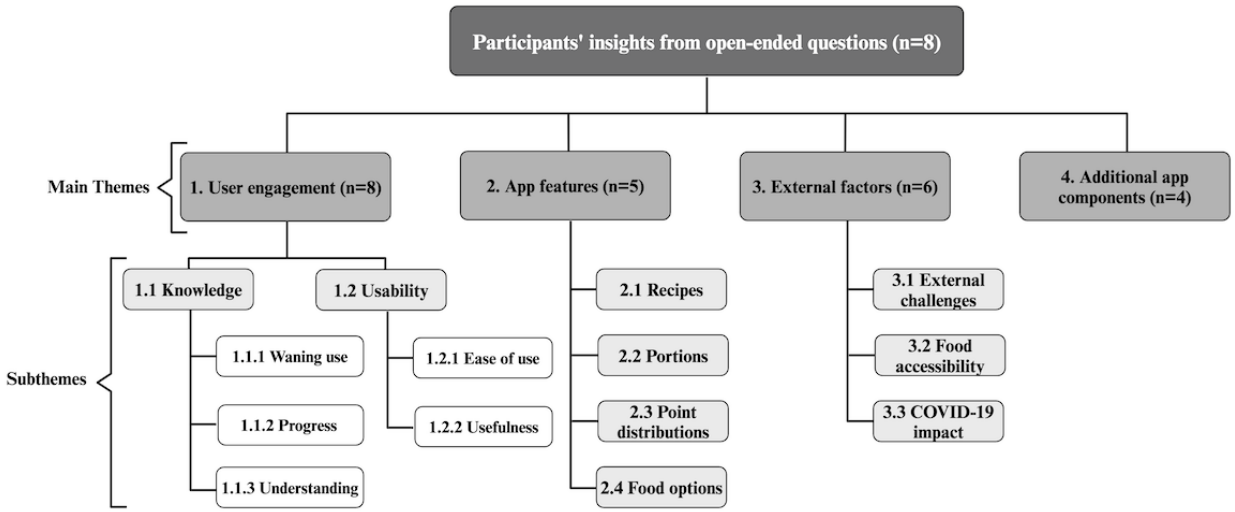
Multimedia Appendix 4 shows the average PortfolioDiet.app use by intervention month. Participants logged into the app an average of 18 (SD 14) days per month over the 12-week intervention period with the number of log-ins trending down over the duration of the intervention but these results were not 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).

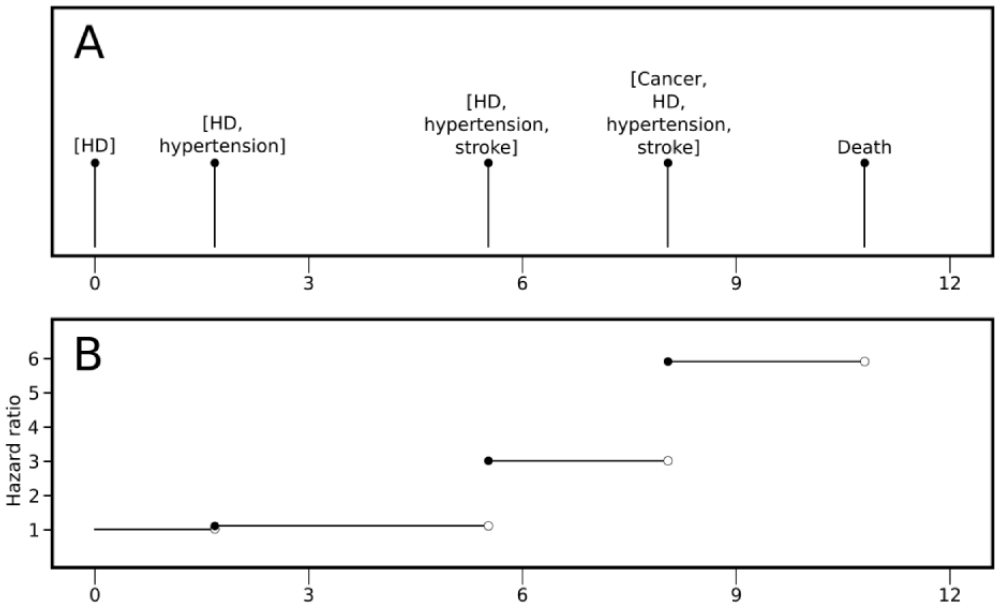
We conducted a model and data-driven selection of interaction effects. Subsequently, we studied the impact on time to death according to the (1) most frequently occurring disease portfolios, (2) most complex disease portfolios in terms of order of interactions, and (3) disease portfolios with the highest 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 ith 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 β_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 β_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 α/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 α.

The selection algorithm runs either until M-way interactions are included or until no k-way interactions are selected in the kth 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 α=.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

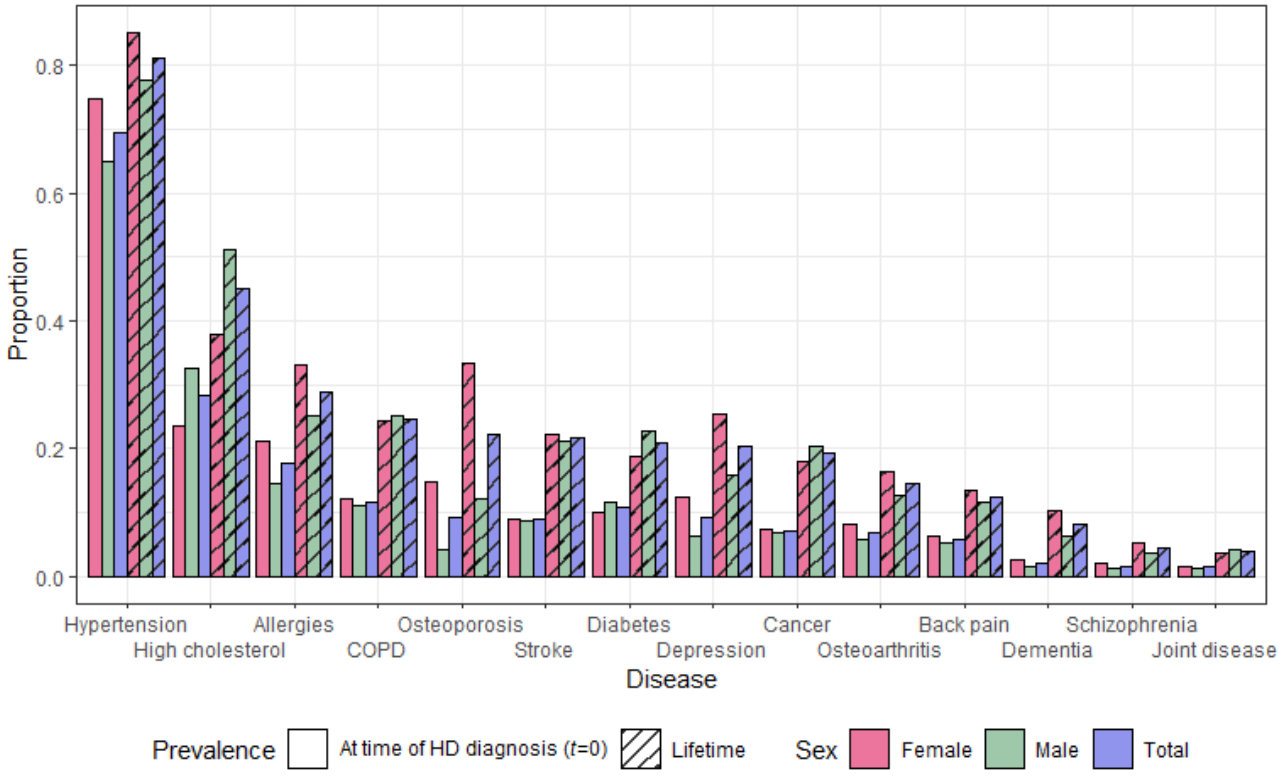
In this study, we used data from the national Danish registries, which are protected by the Danish Data Protection Act, meaning that they can only be accessed after application and subsequent approval. This study did not require additional approval from the Danish Research Ethics Committees or any informed 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.

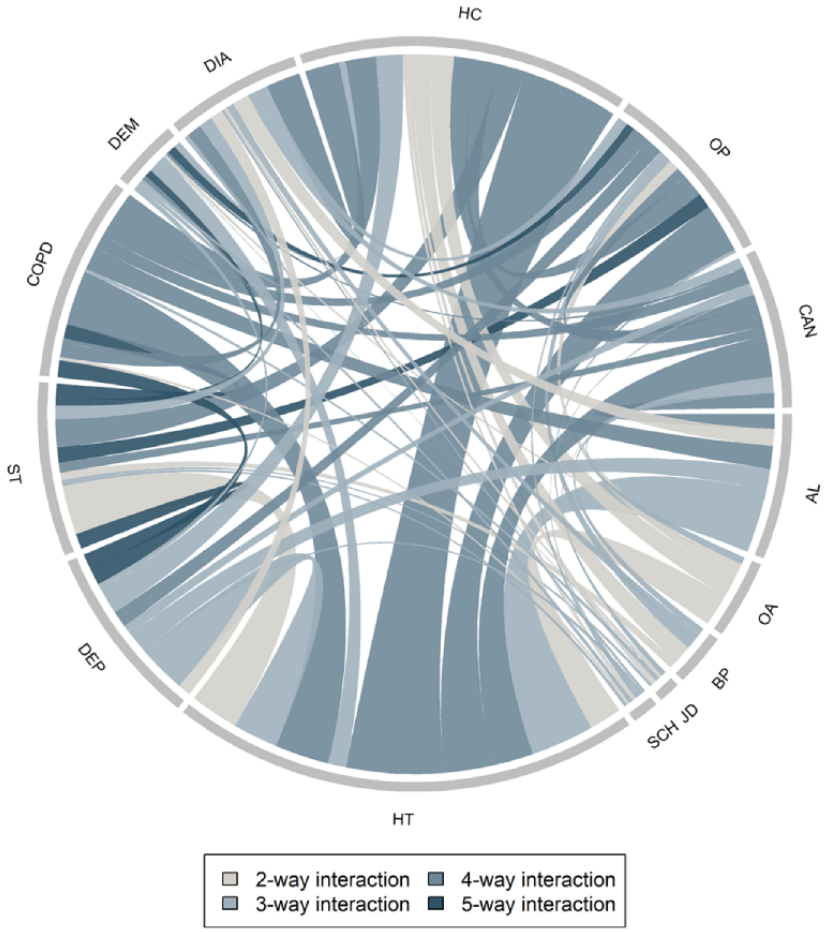


Interactions

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.

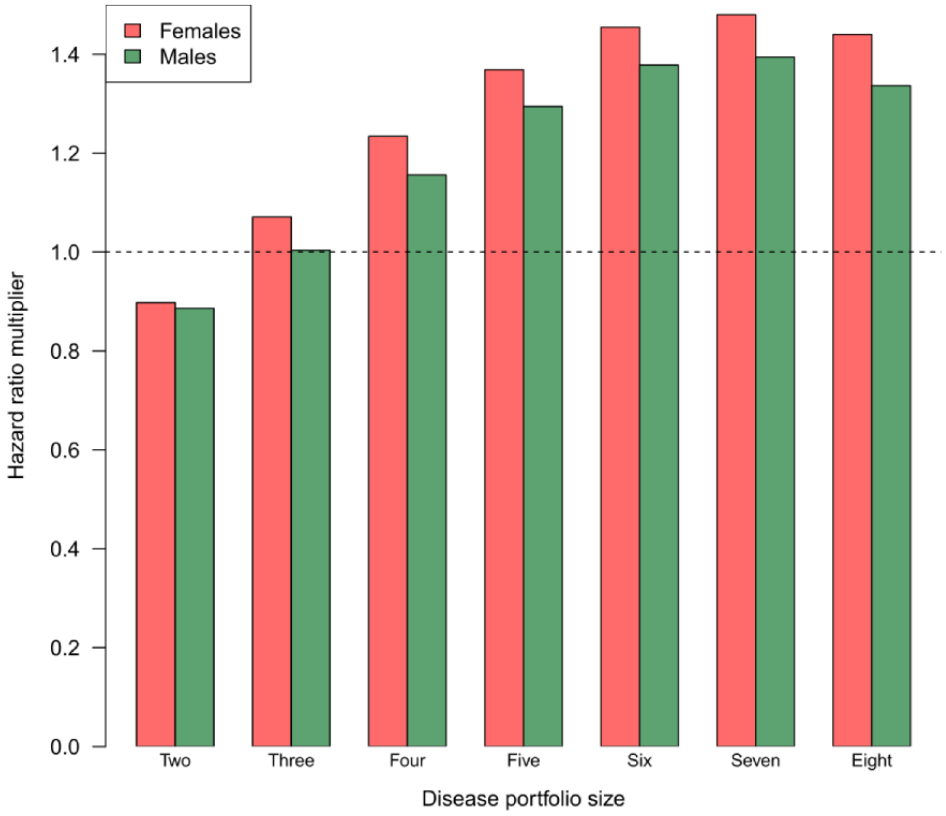


Effects

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.

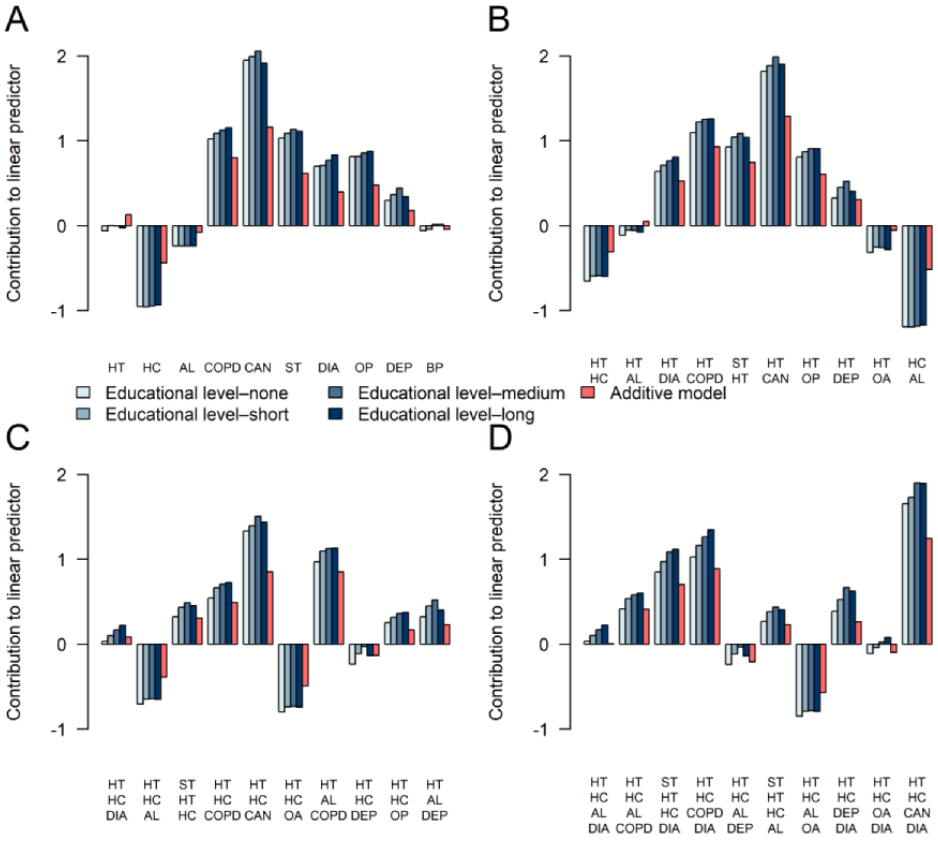
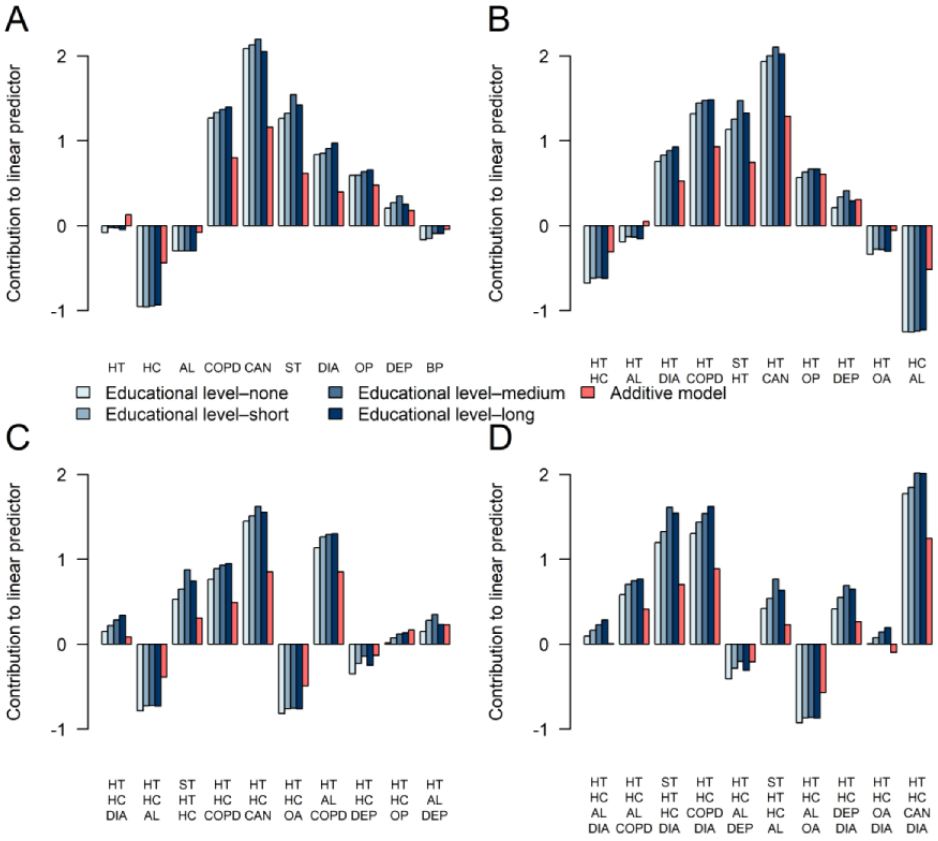


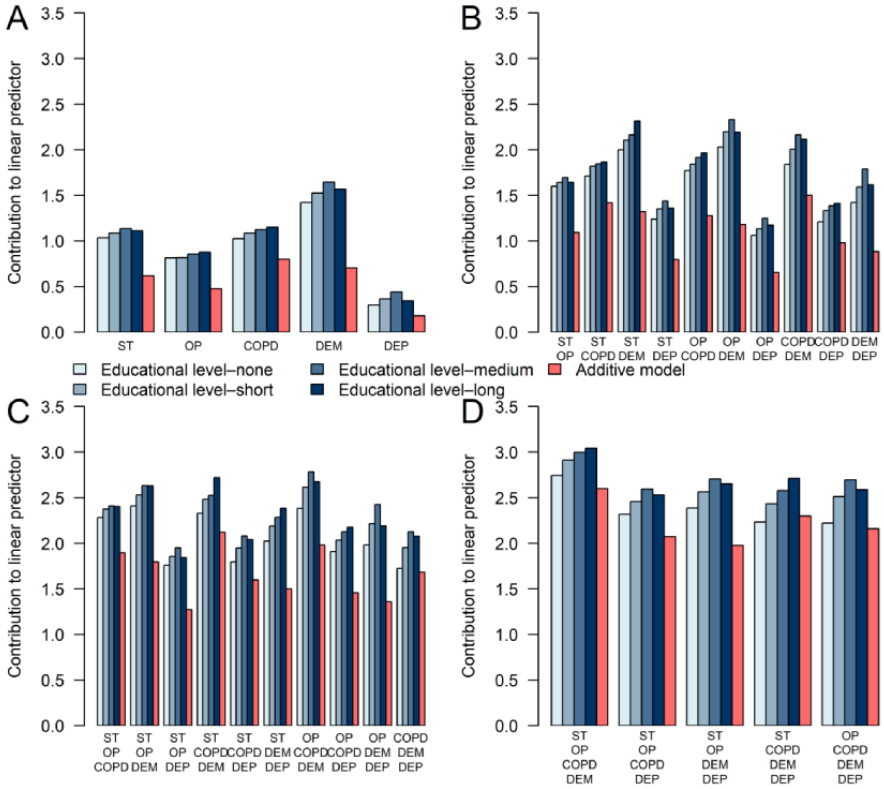
Figure 6. Effects of the 10 most frequent disease portfolio dyads (A), triads (B), tetrads (C), and pentads (D). Effects are shown for female individuals of varying educational attainment levels at the log-hazard rate scale. Comparisons are made to a female individual of the corresponding educational 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.

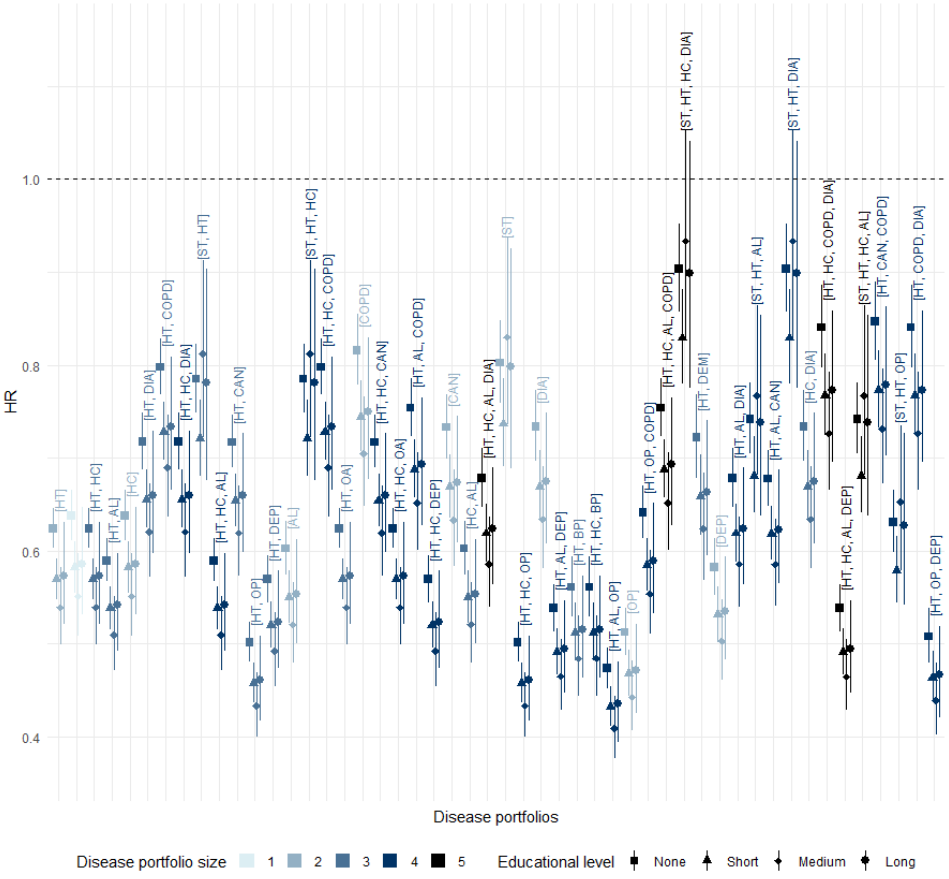
[h]OP: osteoporosis.

[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.

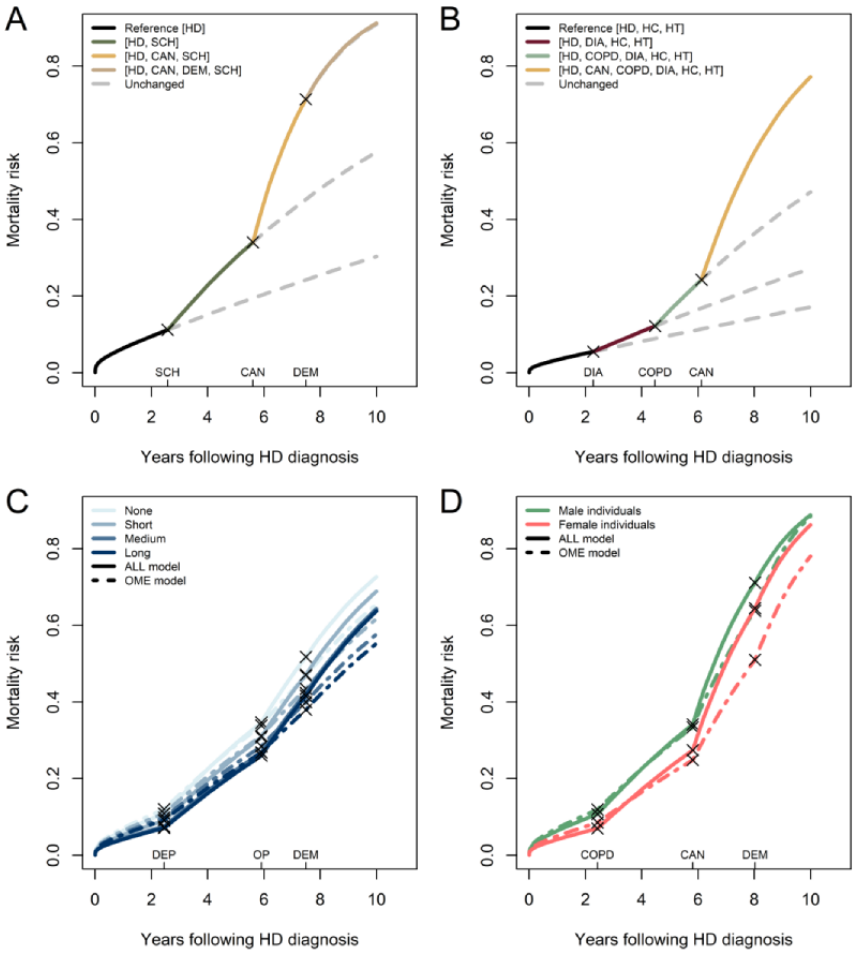
Sex	Disease portfolio size						
	2	3	4	5	6	7	8
HR for male individuals	2.81	2.98	2.74	2.50	2.27	2.08	1.91
HR for female individuals	3.57	3.77	3.43	3.08	2.75	2.45	2.19

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:

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.

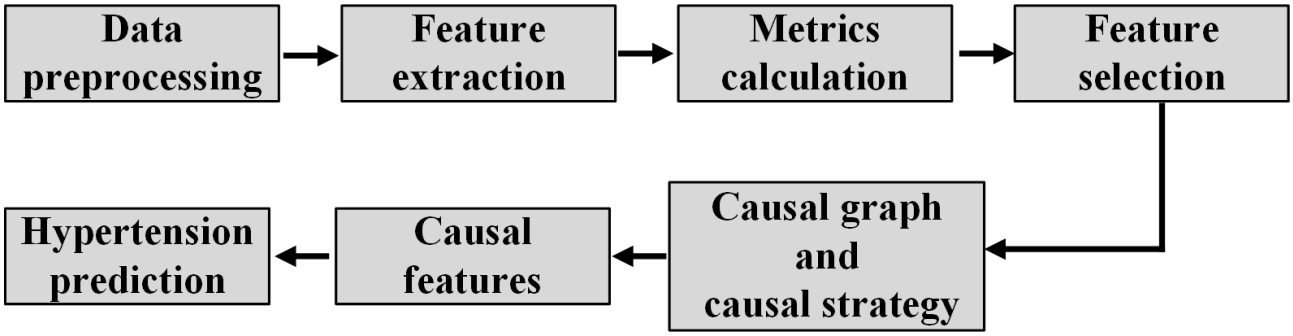
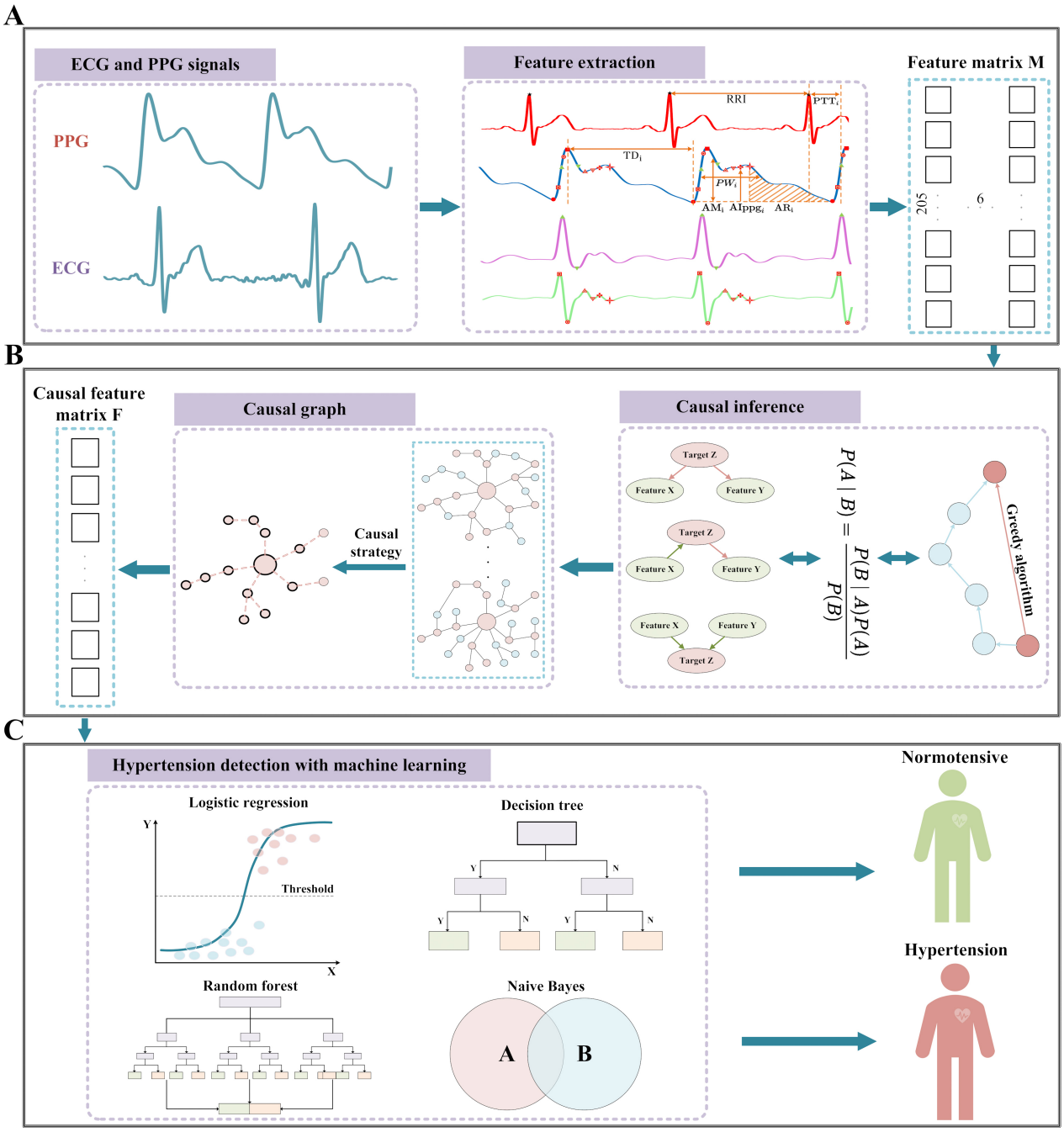


Figure 2. Flowchart of the causal inference for hypertension prediction. (A) Signal preprocessing: 205-dimension beat-by-beat features were extracted from the ECG and PPG as well as the first and second derivatives of the PPG signal (dPPG, sdPPG), and the statistical metrics of these features were calculated 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.

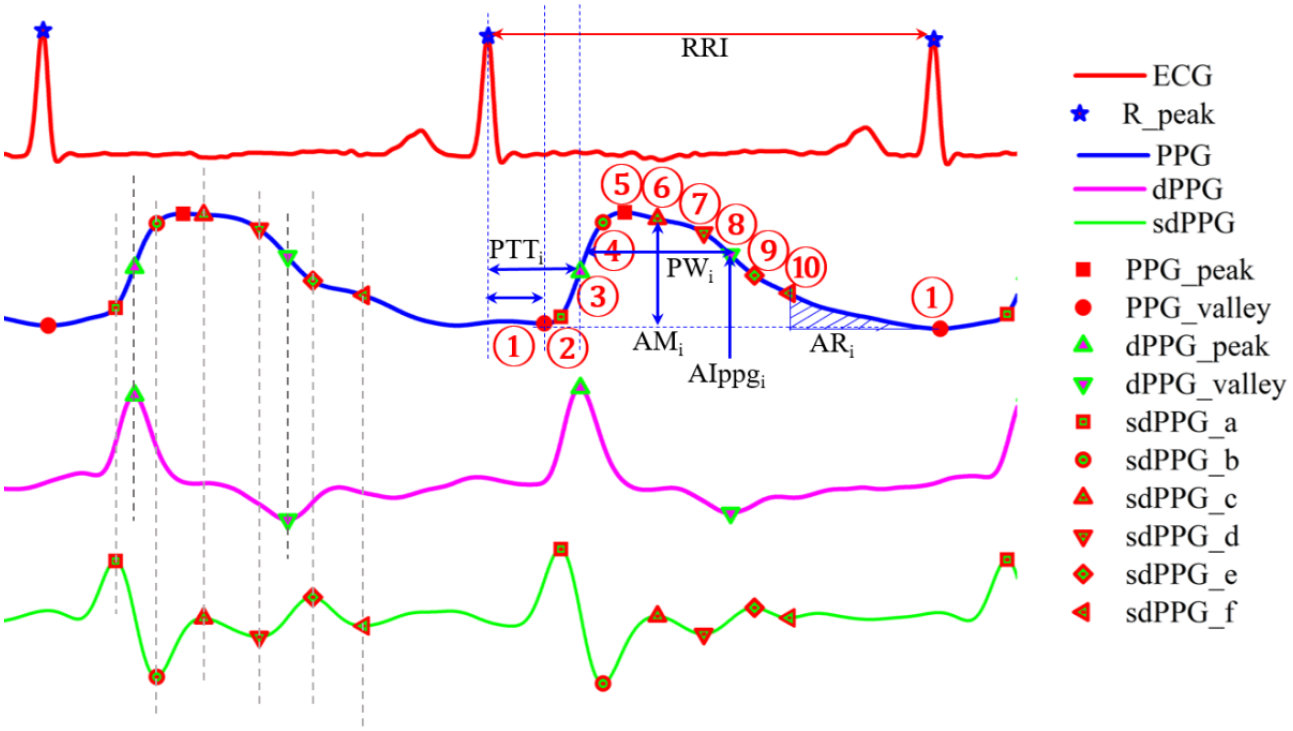
Table . Features extracted from electrocardiogram and photoplethysmogram signals.

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

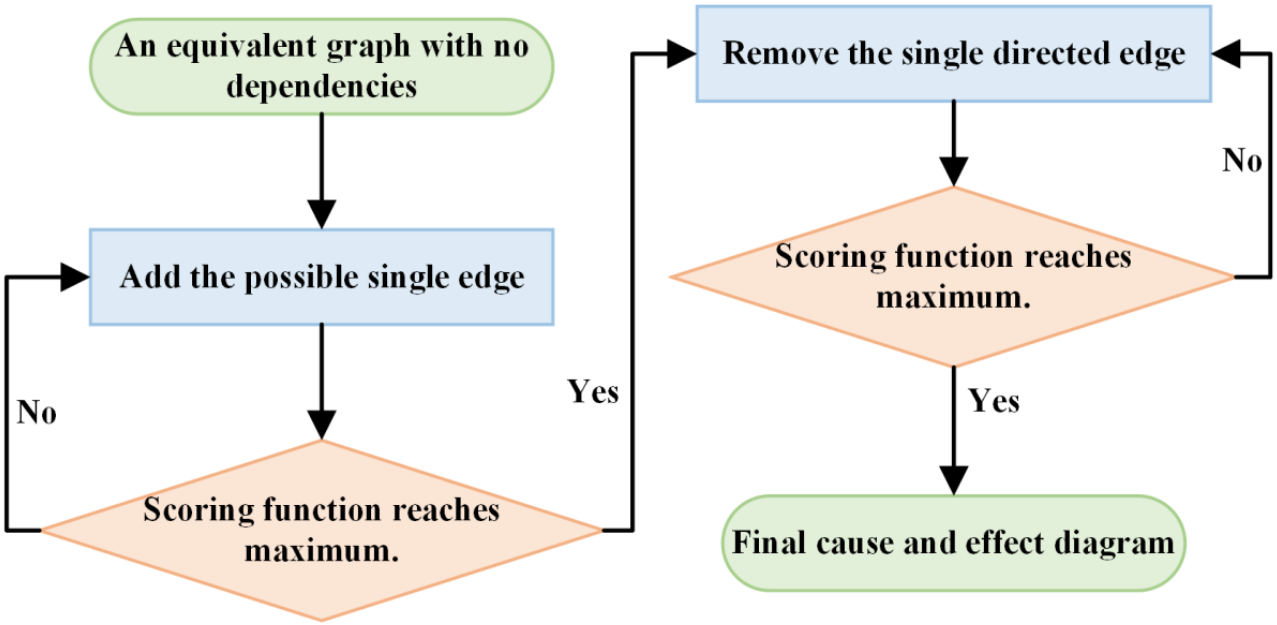
RI: relative rising time, dicrotic diastolic ratio, augmentation index, inflection point area point, slope transit time, ratio of sdPPG (b/a, c/a, (c+d–b)/a, etc), PPG intensity ratio, perfusion index [13].

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.

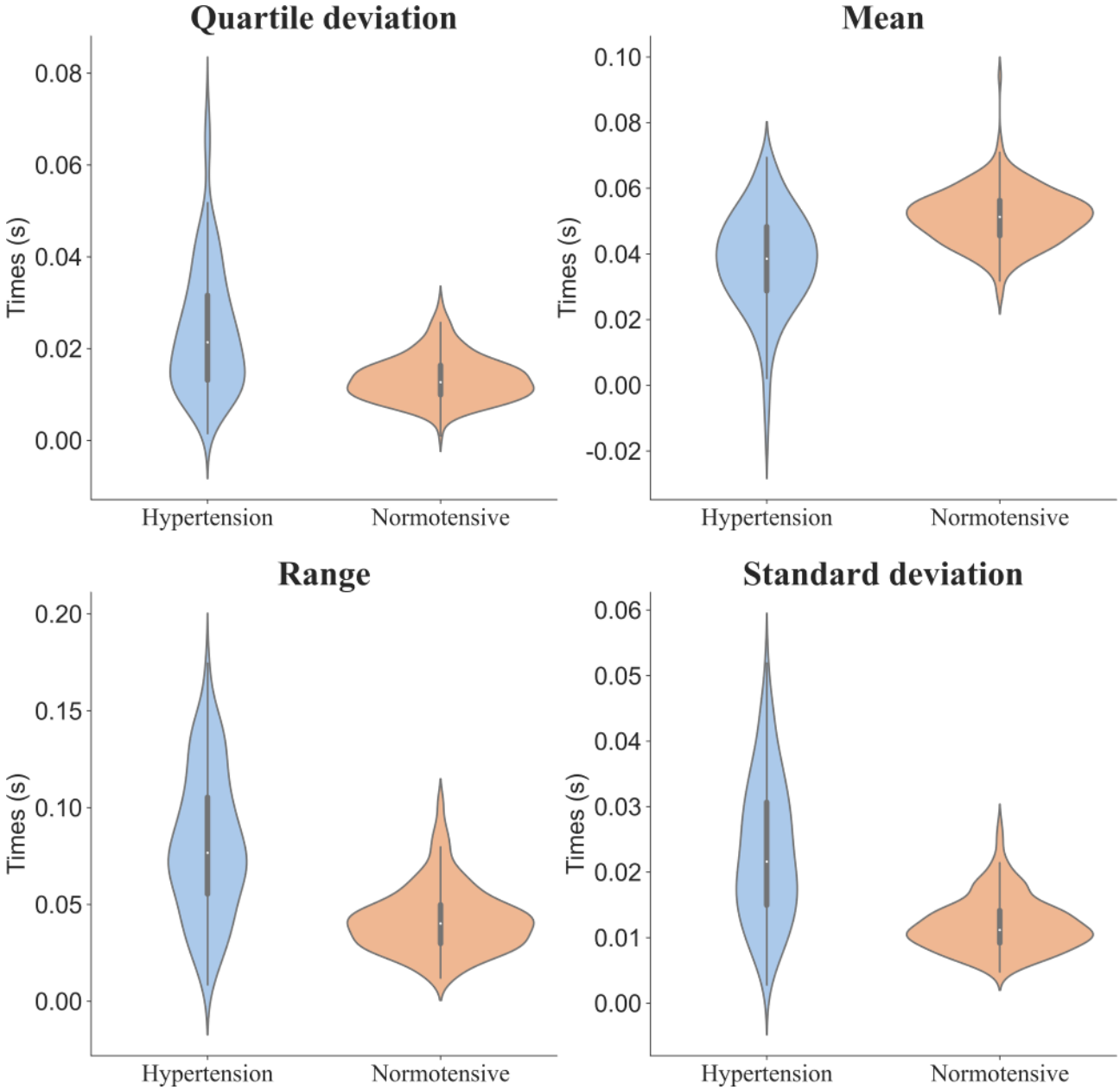
Finally, we needed to organize a feature matrix in which each row represents a sample and each column represents a kind of feature, then input this matrix into the equivalent greedy search algorithm to obtain the causal graph. Prior to this, feature selection is a necessary step to construct the feature 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.

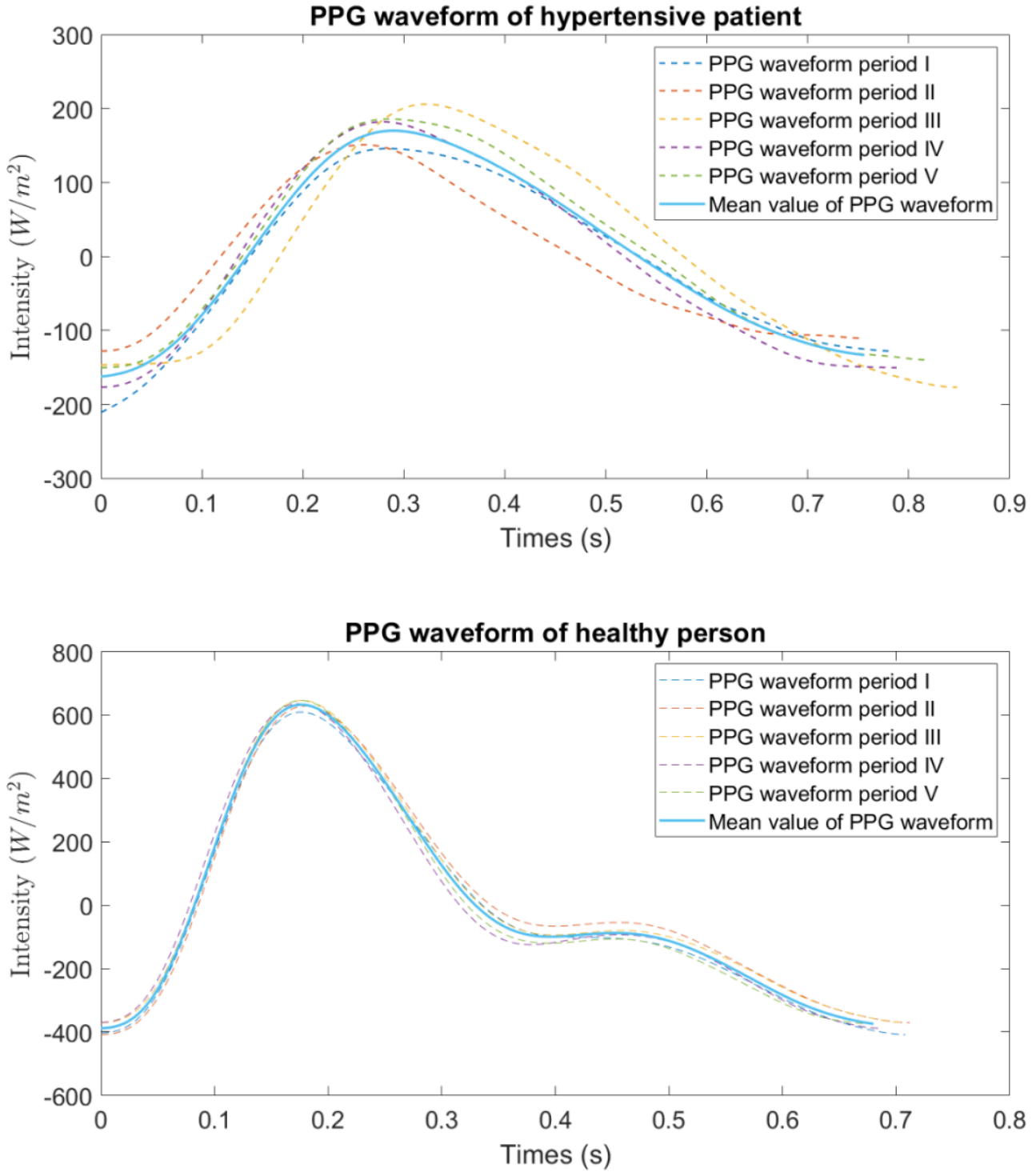
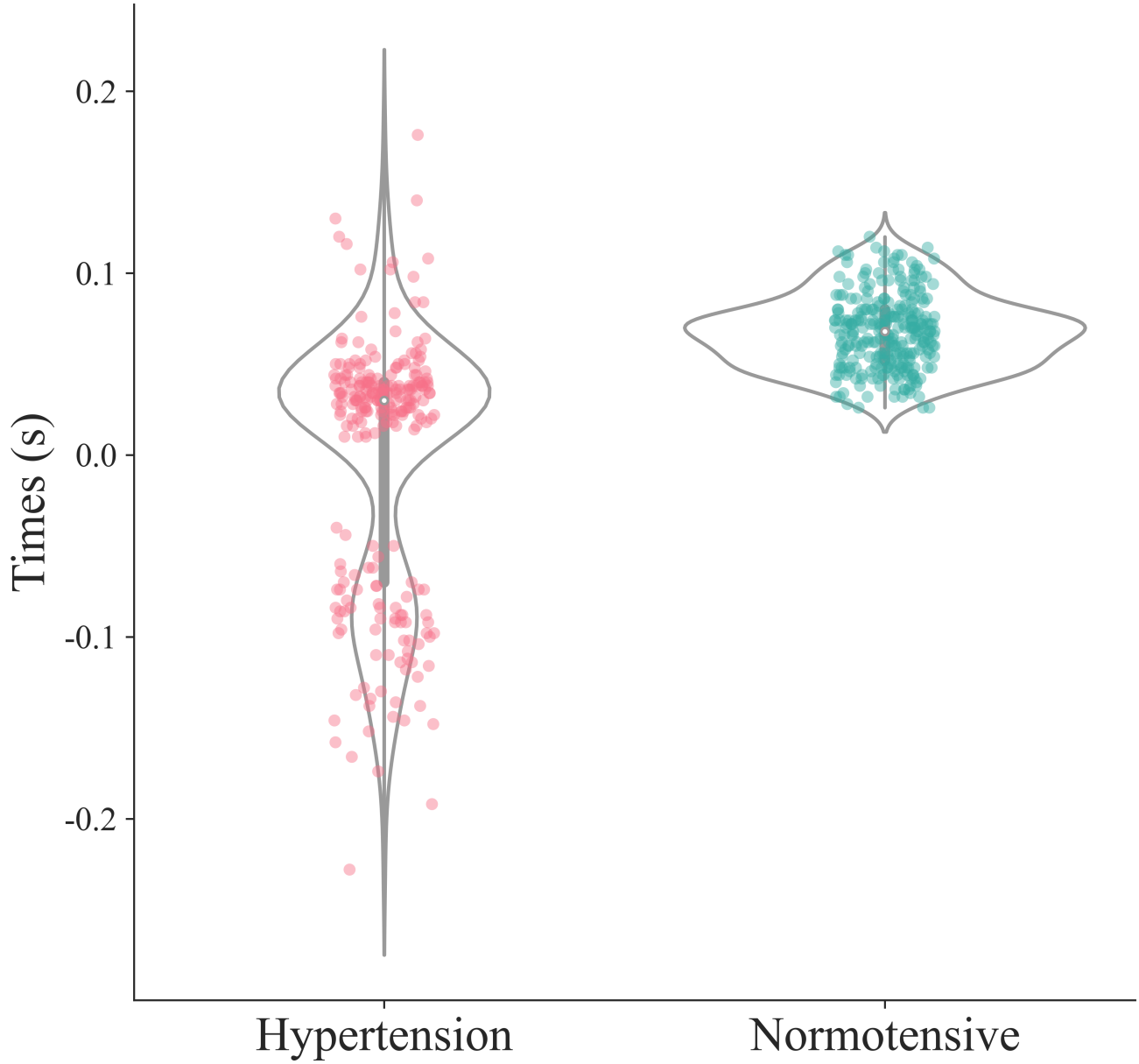


Figure 7. Scatter distribution of feature 52 for normotensive subjects (green) and hypertensive subjects (red).



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.

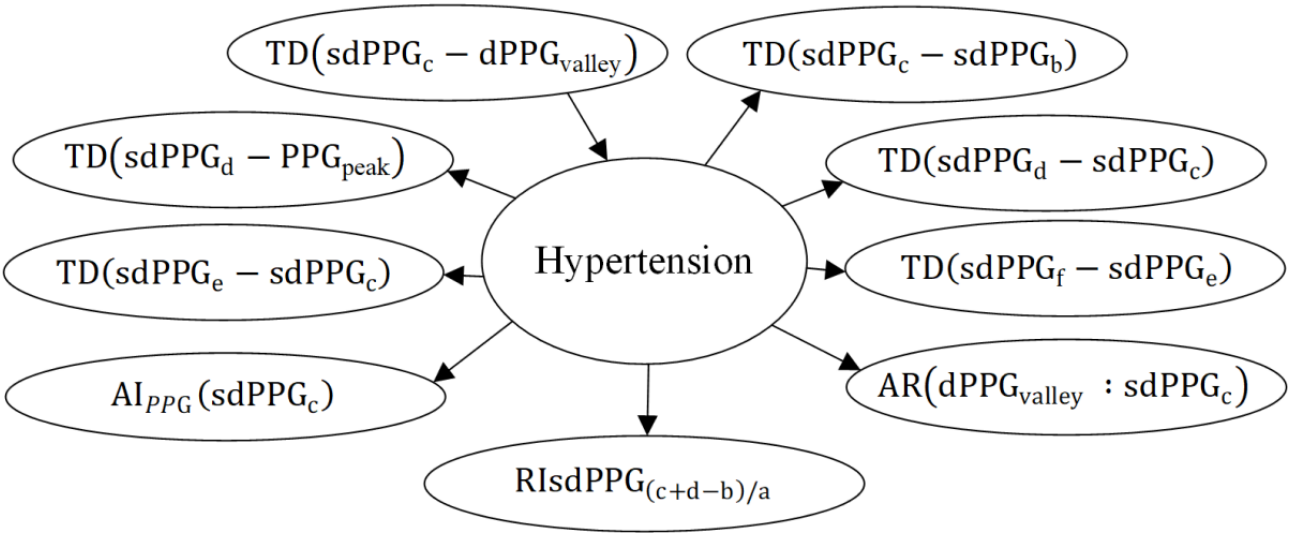


Figure 9. Causal subgraph of hypertension and the features calculated with their range. AM: amplitude; AR: area under the PPG curve; dPPG: the first derivative of PPG; P-R: precision-recall; PPG: photoplethysmogram; RI: physiological meaningful relative index; sdPPG: the second derivative of PPG; TD: time duration.

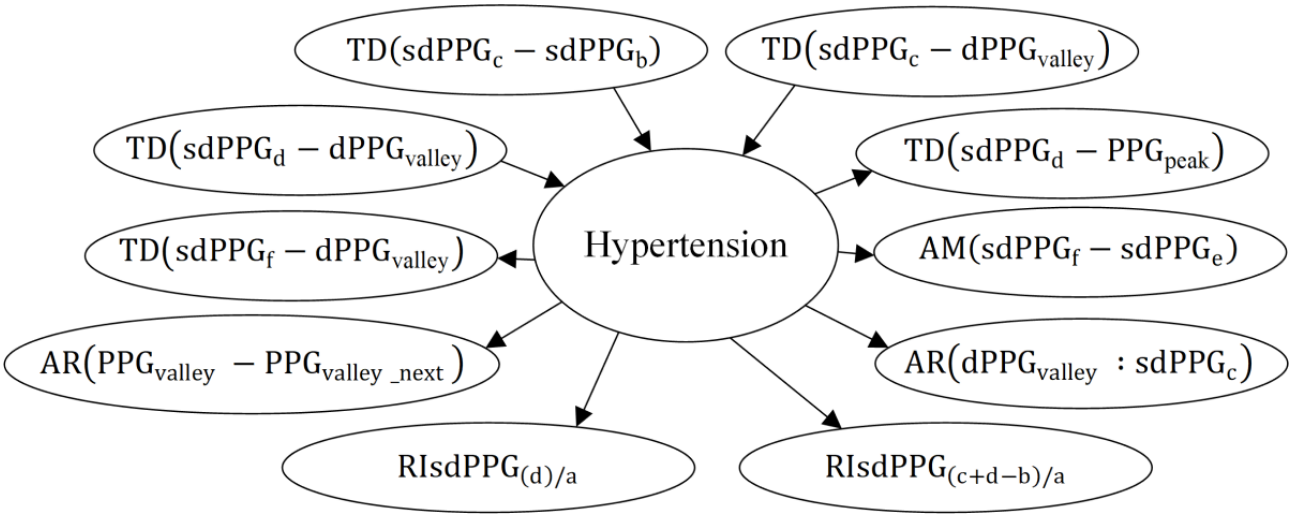
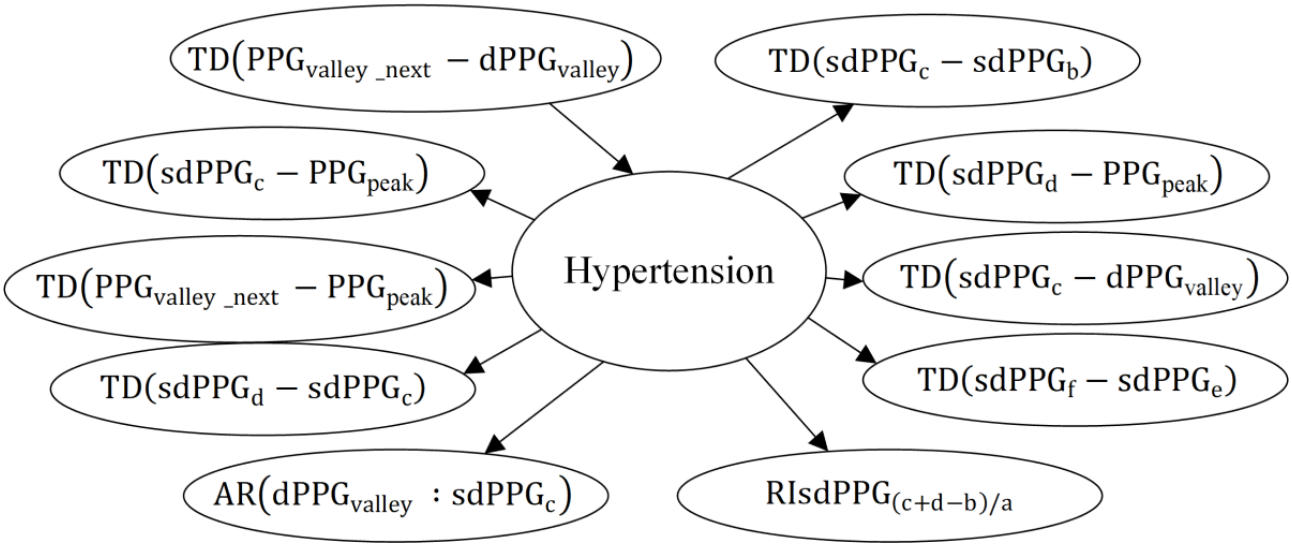


Figure 10. Final causal graph. AR: area under the PPG curve; dPPG: the first derivative of PPG; P-R: precision-recall; PPG: photoplethysmogram; RI: physiological meaningful relative index; sdPPG: the second derivative of PPG; TD: time duration.



Hypertension Classification Results

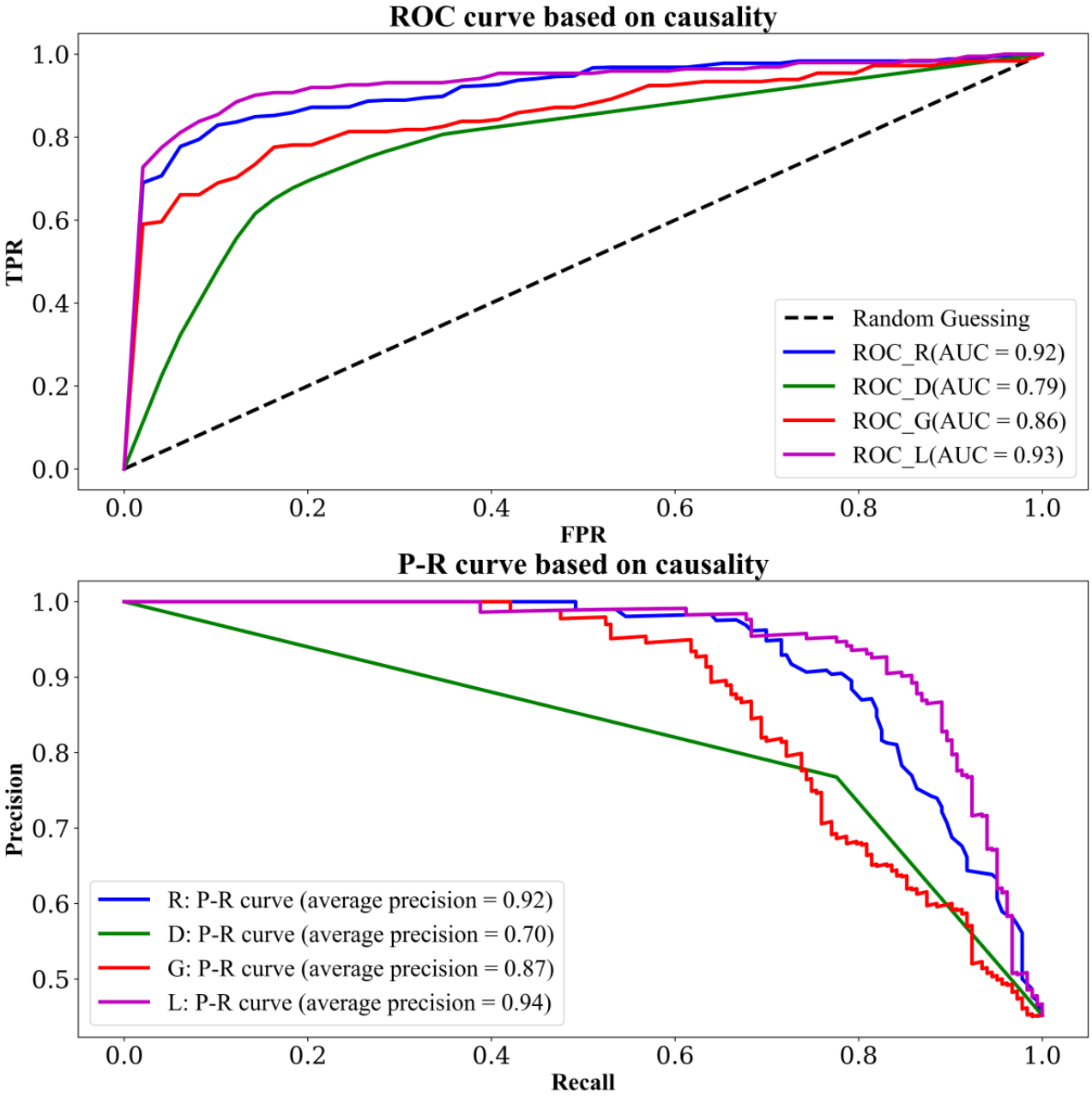
In this subsection, we used multiple classifier algorithms for hypertension classification prediction. First, we primarily utilized logistic regression and other classification algorithms based on causal feature variables for hypertension classification. The classification performance is presented in 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.

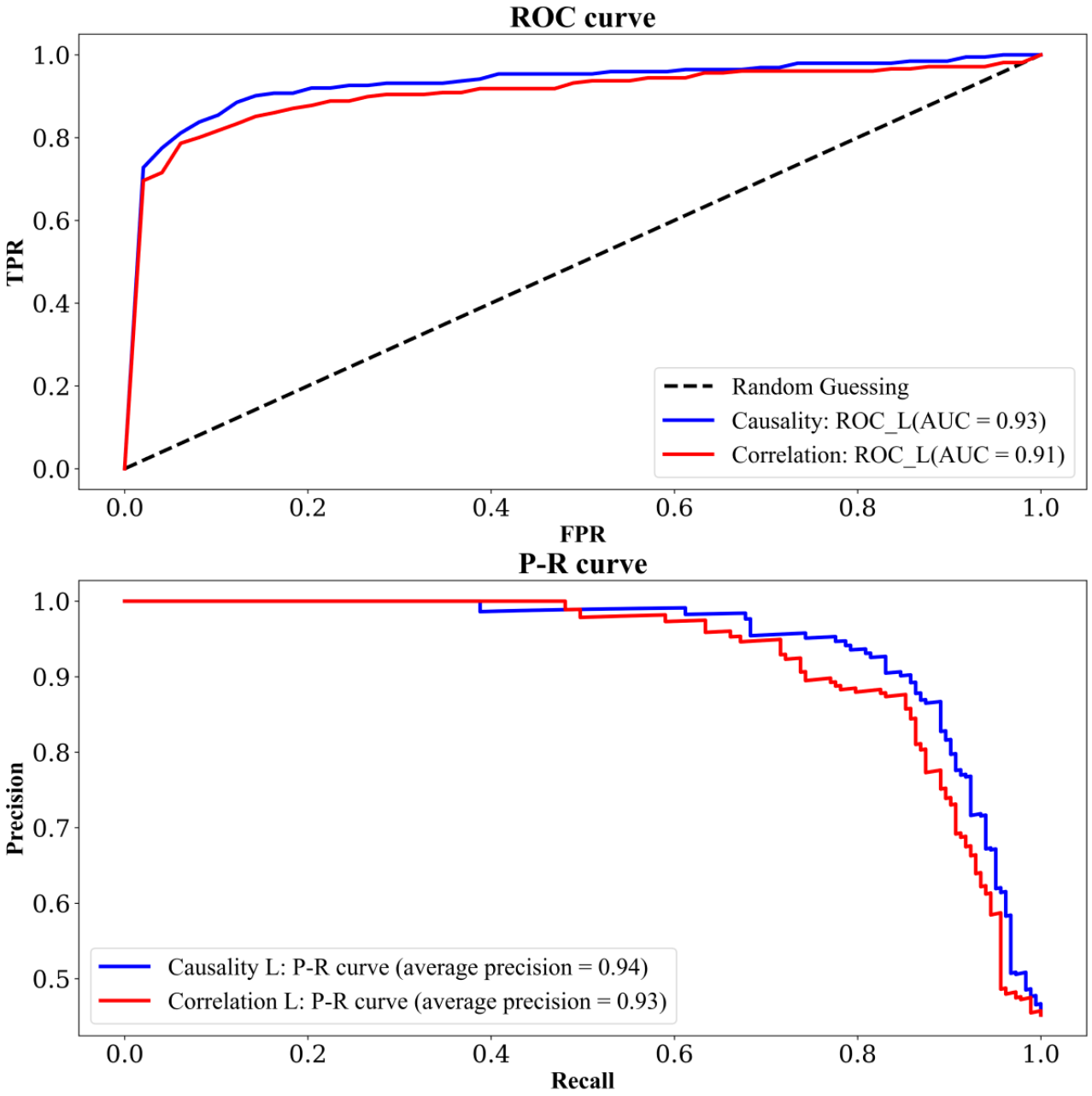
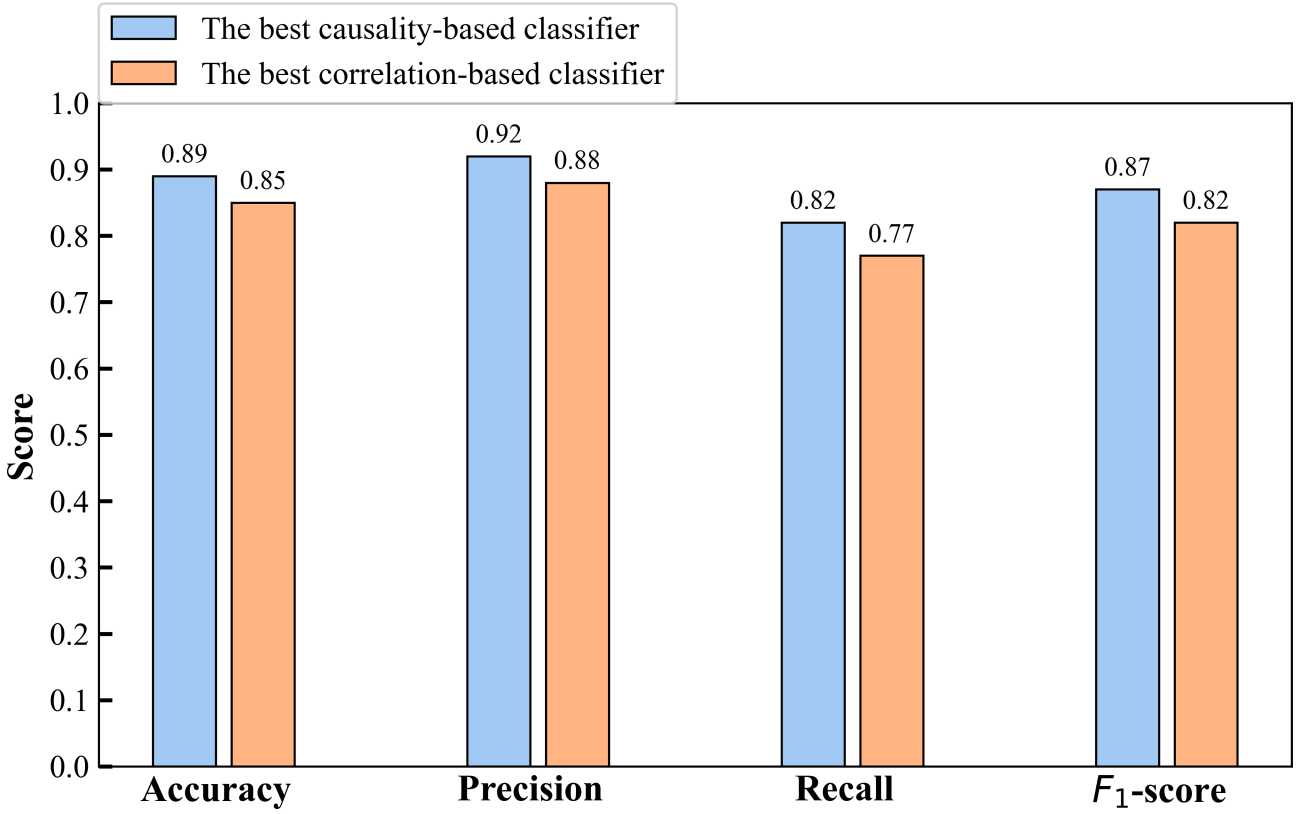


Figure 13. Histogram of evaluation metrics for the best classifier of causality and correlation.



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

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.

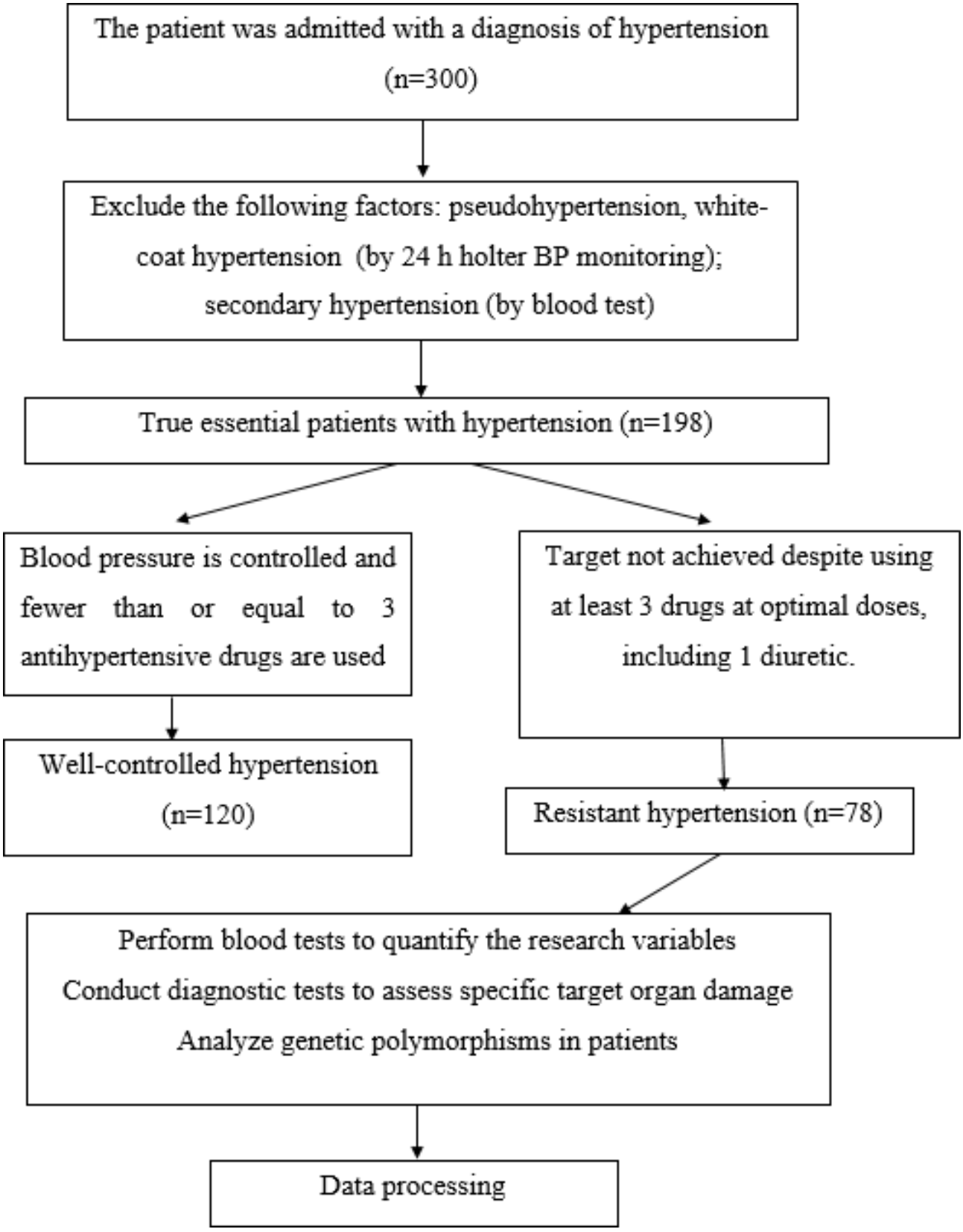


Table . Clinical characteristics of patients with RH[a].

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

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.

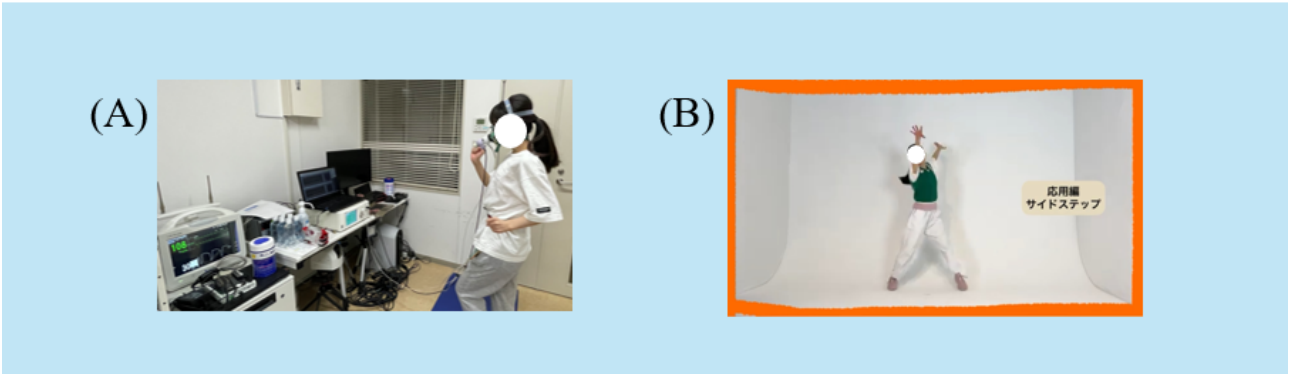
We included outpatients with high BP from the Juntendo University Department of General Medicine. These patients with hypertension had been diagnosed with hypertension and were receiving regular oral medication. The patient was invited to participate in this study by their primary physician, whom they 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).

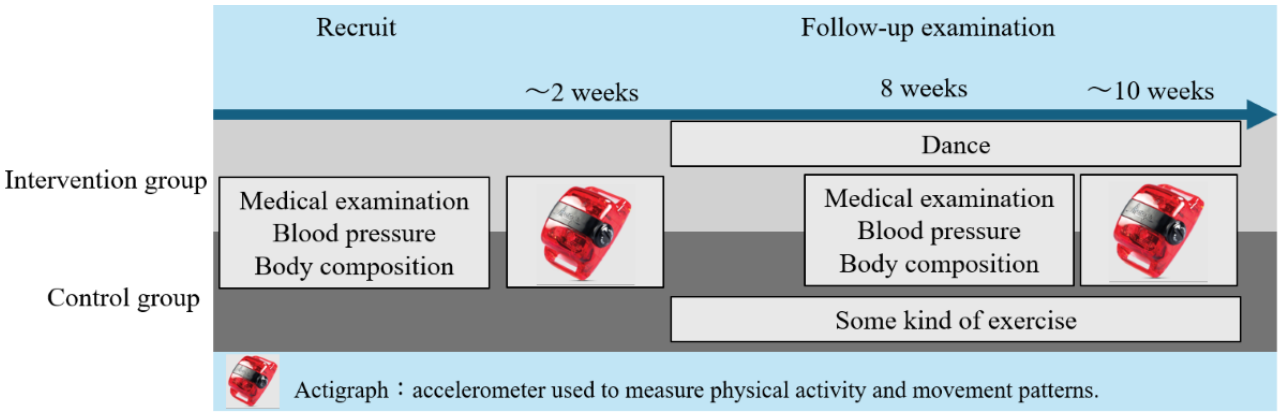
Participants from both groups were instructed not to change their lifestyle 2 weeks from the day of recruitment and to wear an ActiGraph continuously during this period, except during sleep and bathing. ActiGraph is a 3-axis accelerometer (wGT3X-BT ActiGraph, ActiGraph, LLC). Actigraph triaxial accelerometers 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.

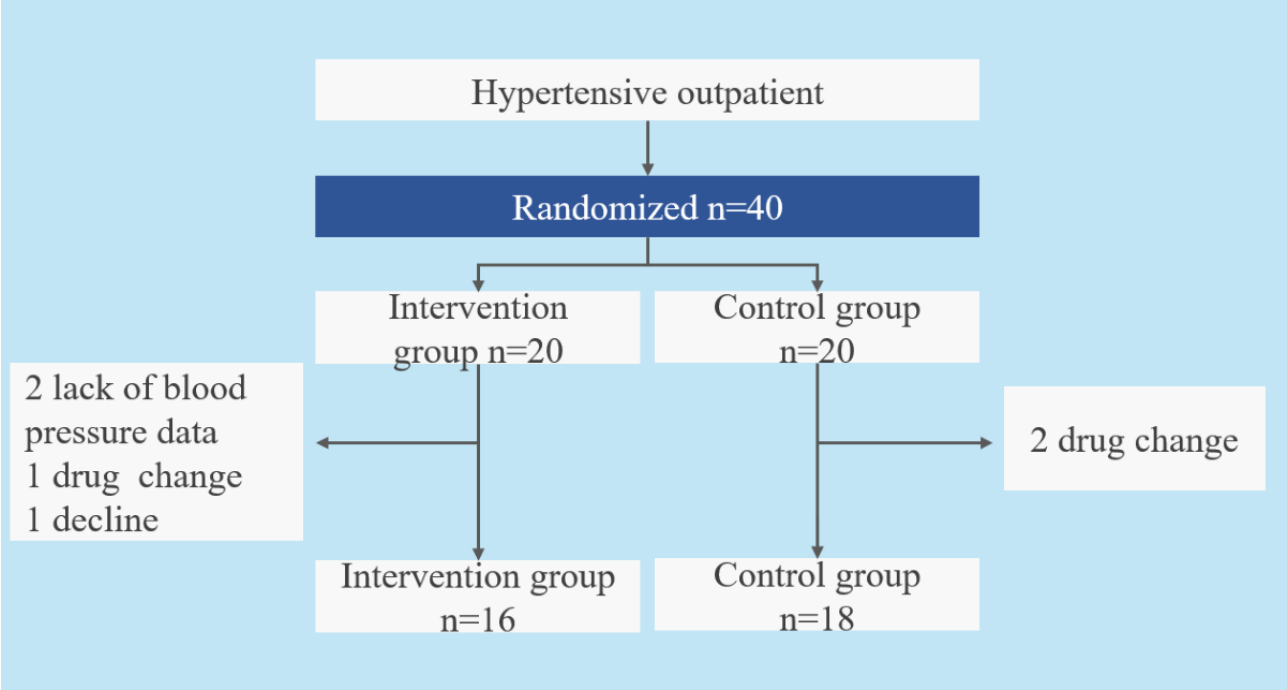


Table 1. Characteristics comparing intervention and control groups[a].

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.

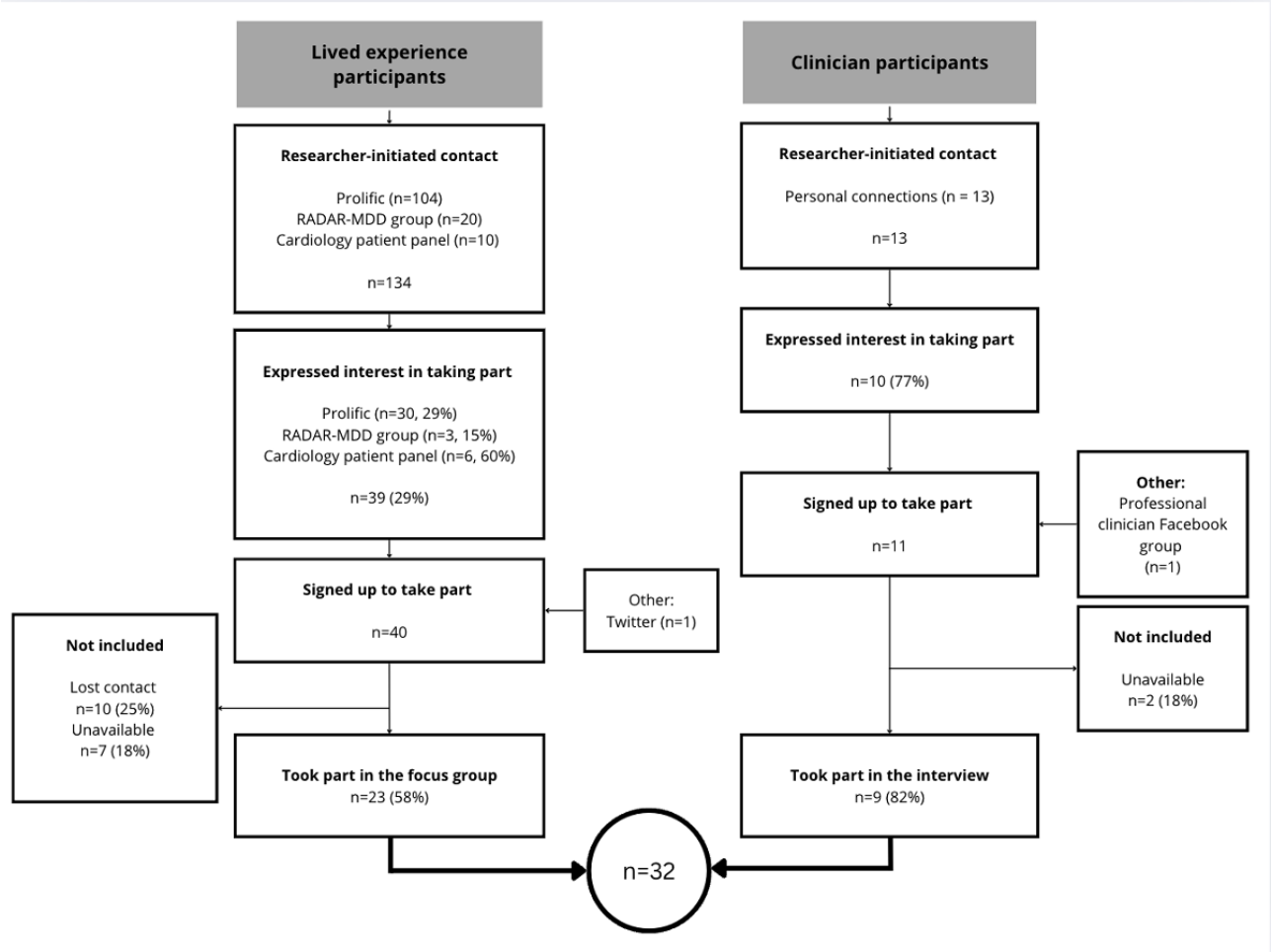


Table 1. Demographic characteristics of the sample (n=32).

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:

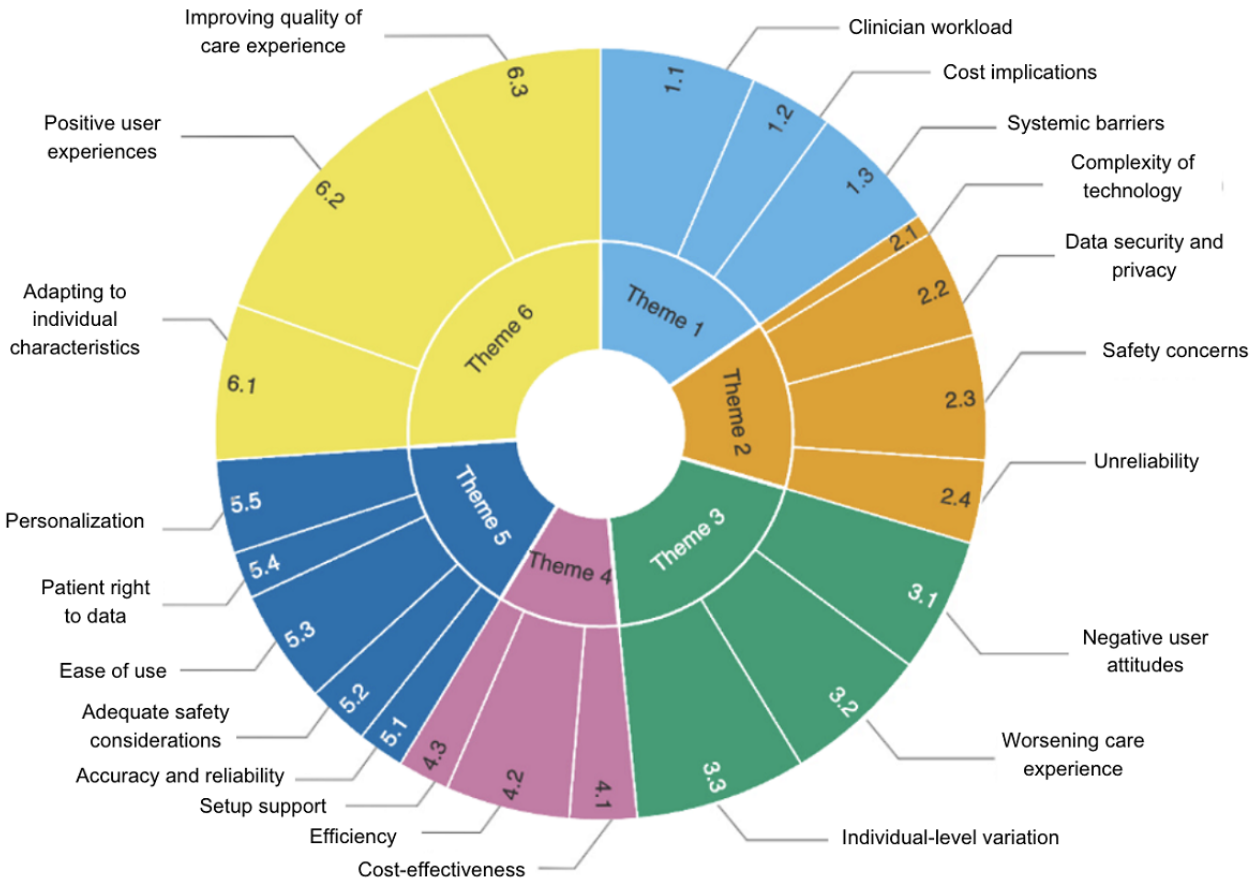
> *It just gives you peace of mind. And obviously with your family members. They put the knowledge around them as well...So that's it's a no brainer really. It's got to help.* [LE8; male; aged 61 years]

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

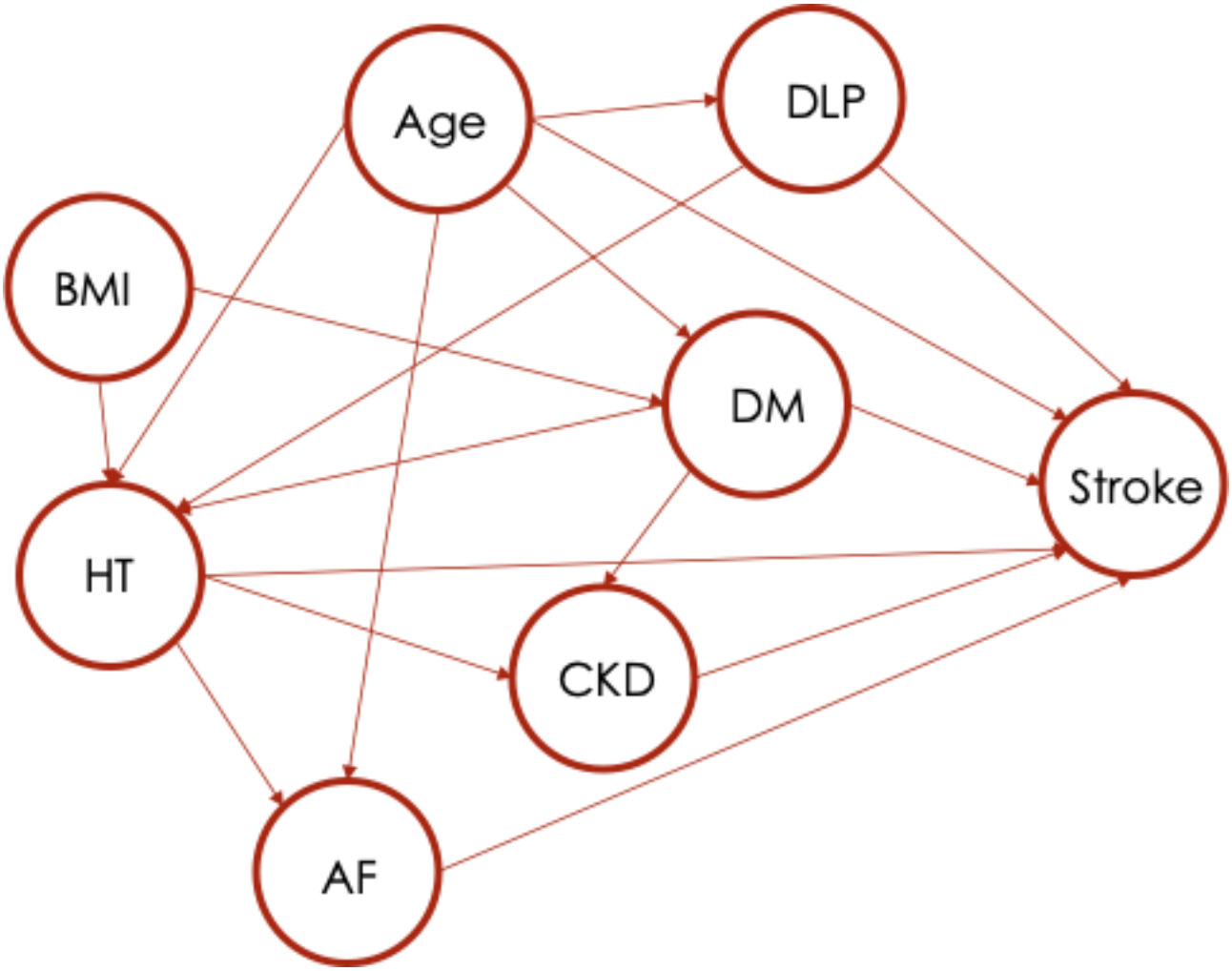
Overview

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

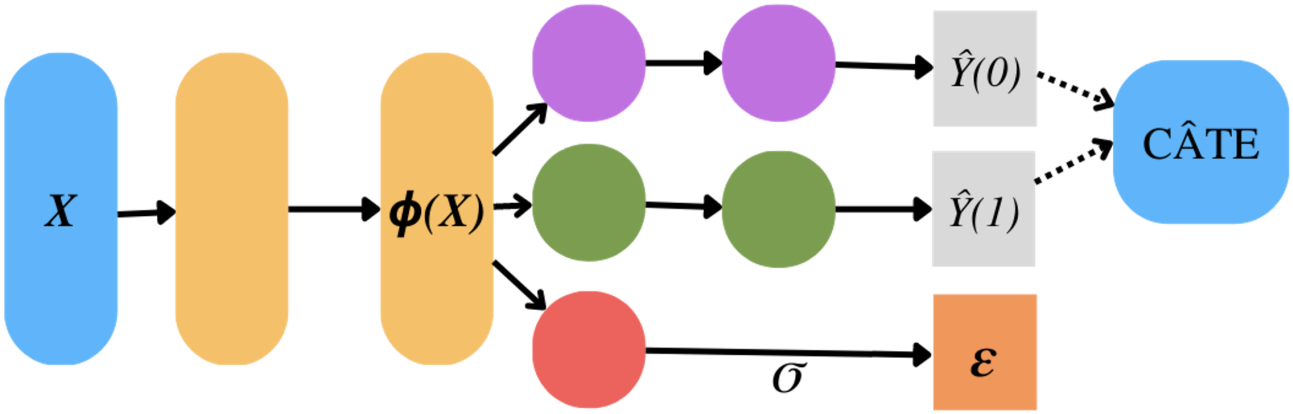
model–based inferences from the DoWhy library were used for medication analysis to quantify the causal effects of direct and indirect pathways, termed natural direct effect (NDE) and natural indirect effect (NIE), respectively [4,26]. NDE (Y1,M(0)x−Y0,M(0)x) refers to the change in the outcome of an 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 $(X) \in$, 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.

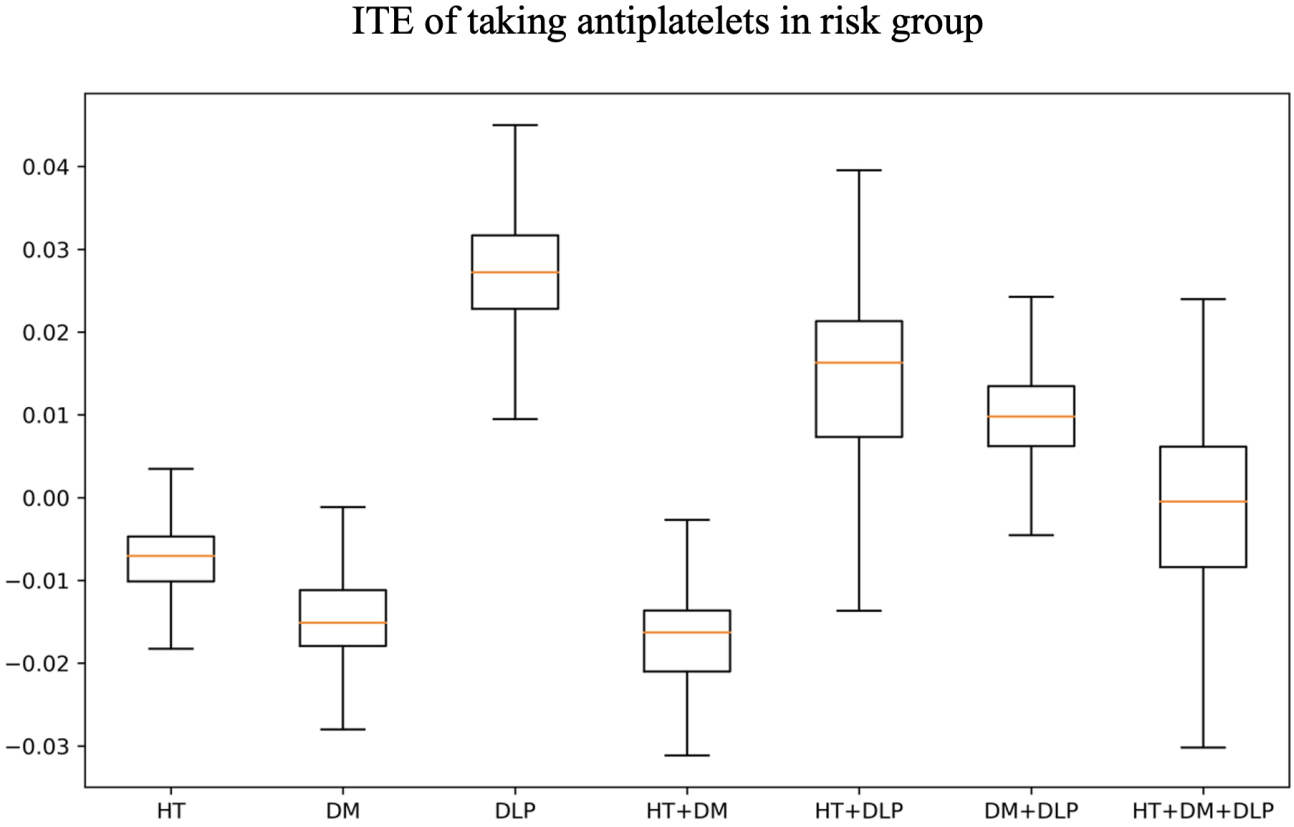
In the context of DML, the nonparametric model estimates were slightly smaller than those for the linear model, with risks of 0.086 (95% CI 0.0849-0.0871), 0.015 (95% CI 0.0145-0.0155), 0.006 (95% CI 0.0057-0.0063), and 0.0 (95% CI −0.0001 to 0.001) for AF, HT, DM, and DLP, respectively, whereas the 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

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.

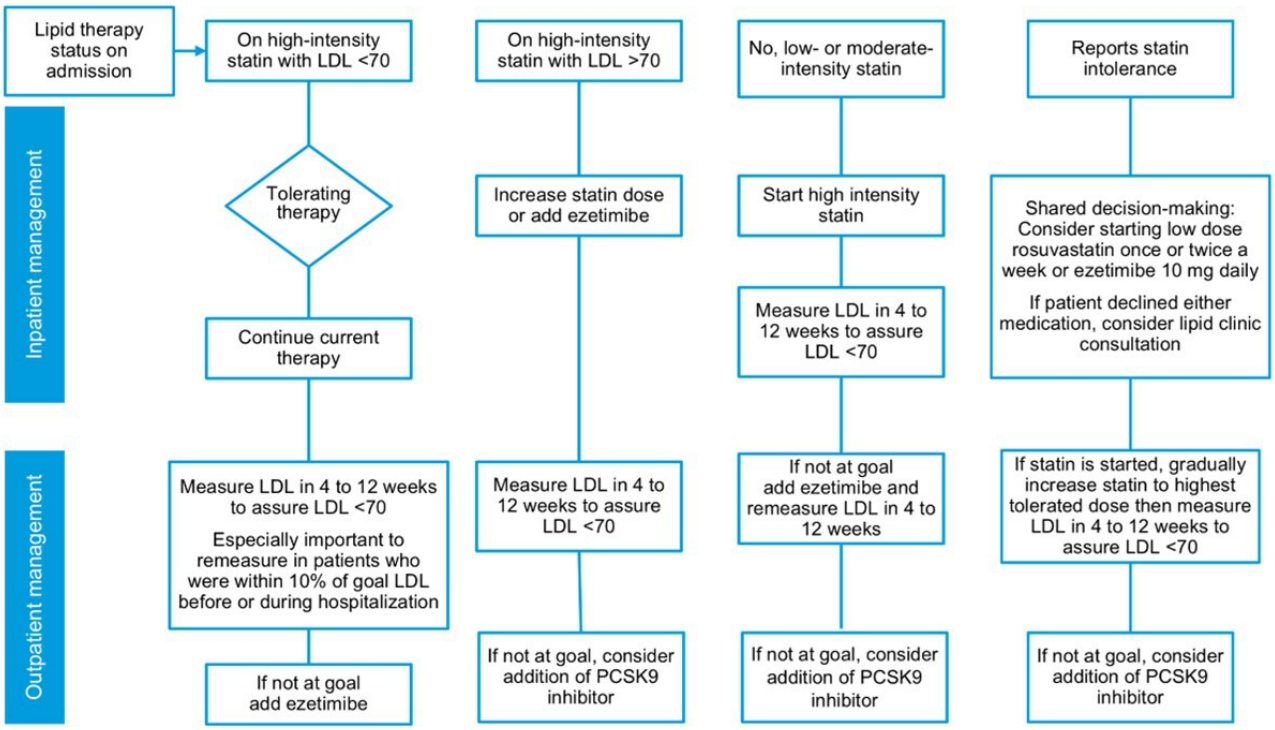
The cardiology pharmacist group consisted of 9 pharmacists who received training and instructions regarding implementation of the intervention in the form of presentations at staff meetings and written documents shared via emails describing project goals and pharmacist roles. At the beginning of each 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 χ^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 χ^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

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.

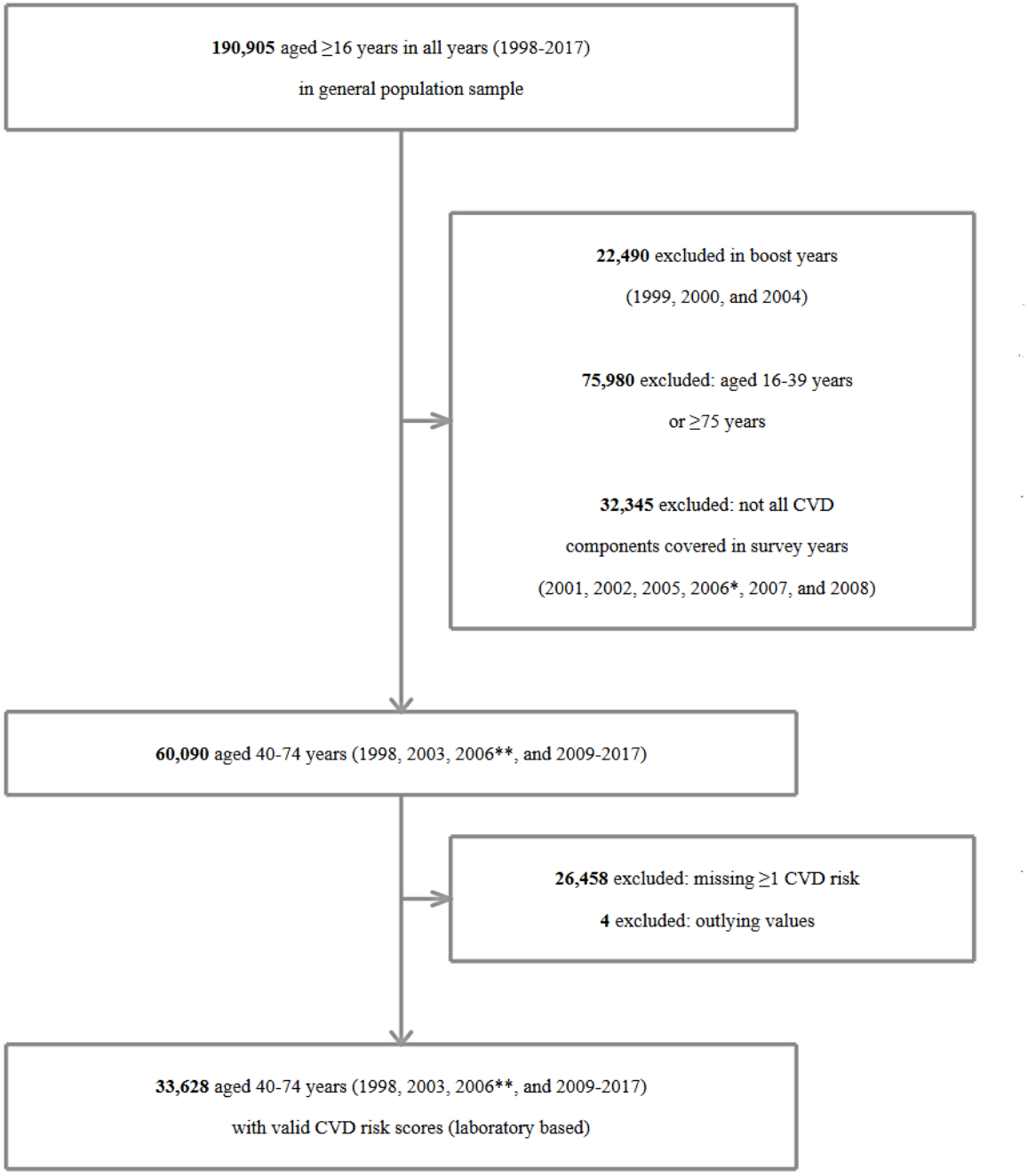
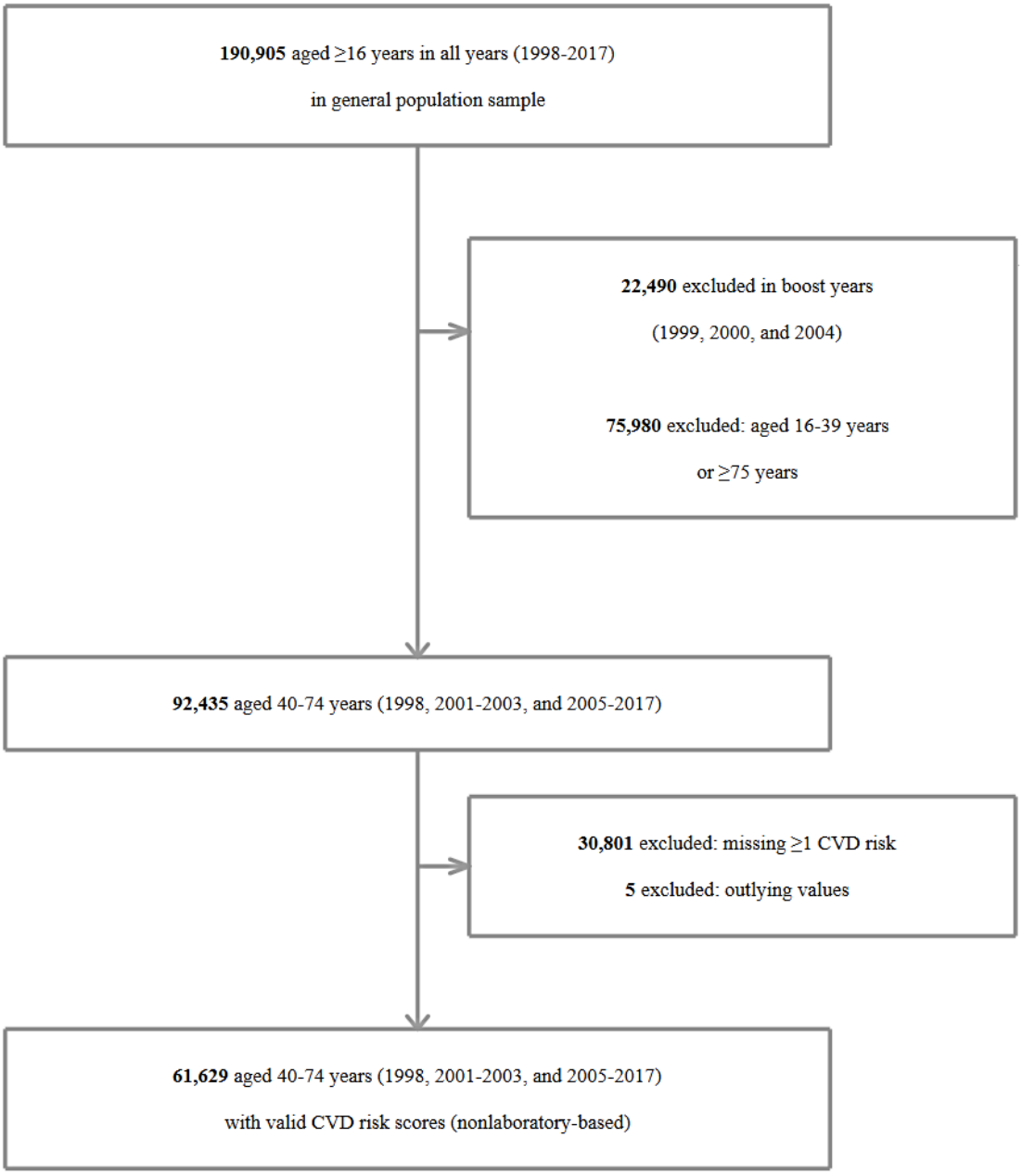


Figure 2. Flowchart of participants included in the estimation of changes over time in cardiovascular disease (CVD) risk (non–laboratory-based scores).



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.

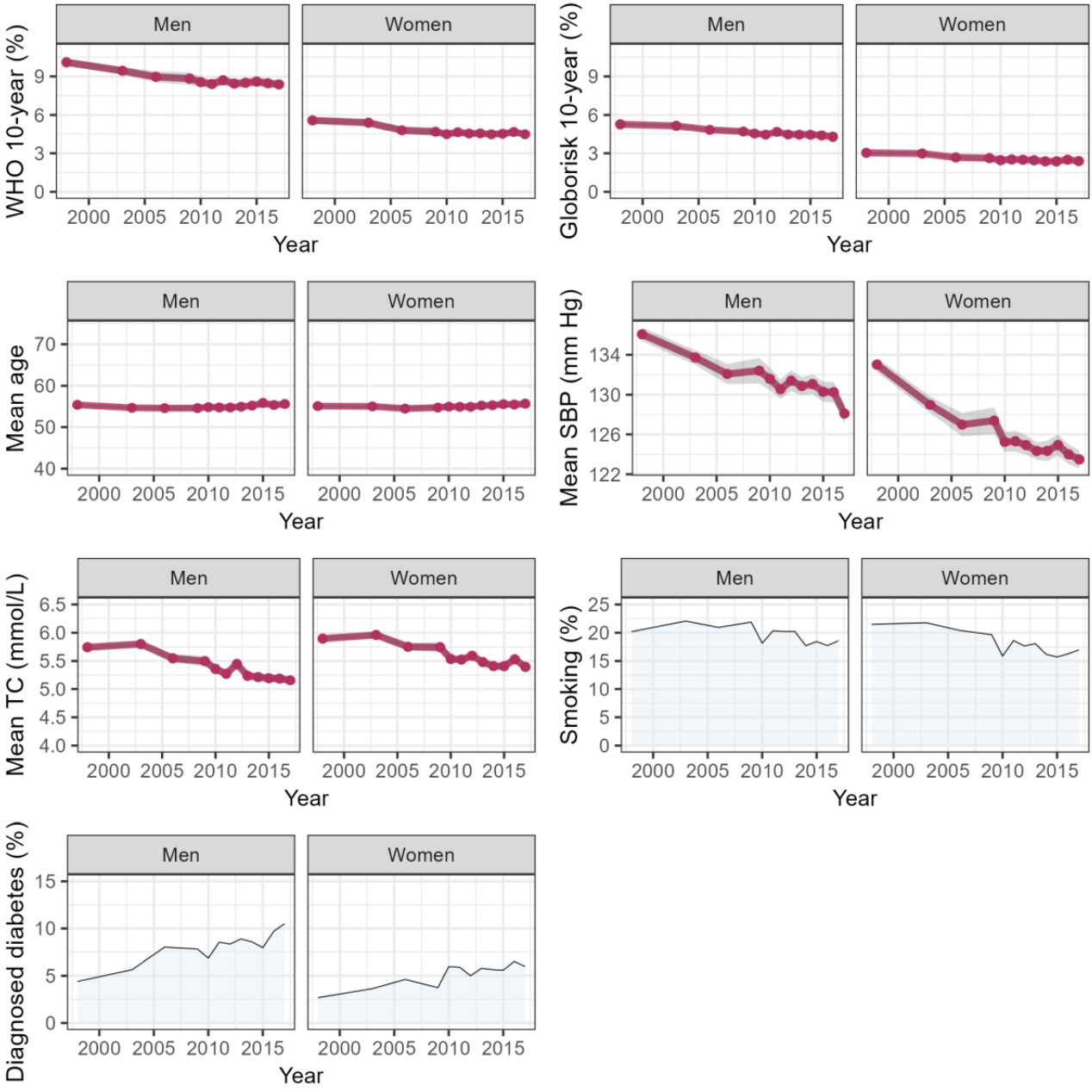
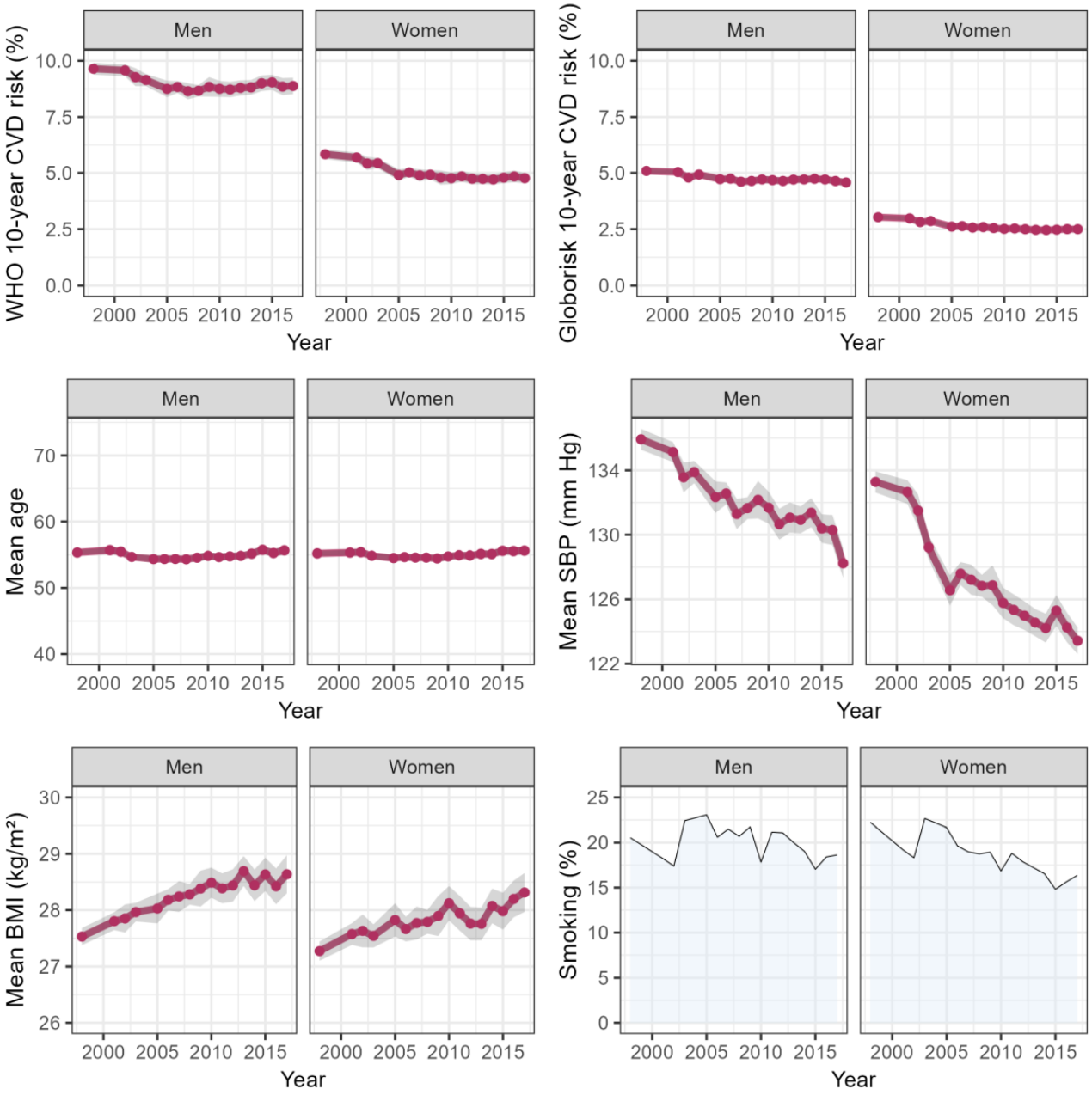


Figure 4. A 10-year cardiovascular disease (CVD) risk score (non–laboratory-based) and its components by survey year and sex. SBP: systolic blood pressure; WHO: World Health Organization.



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:

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.

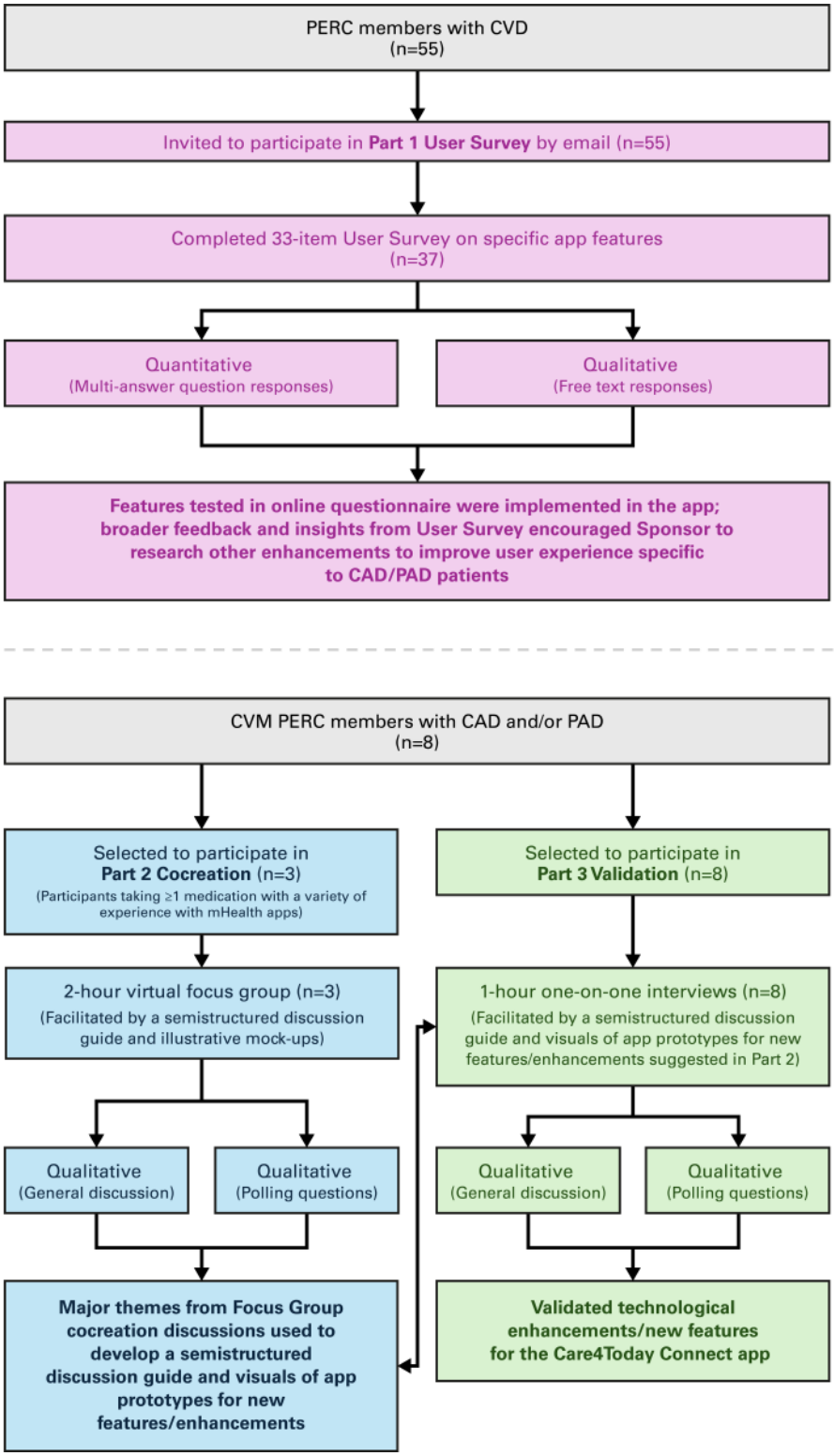


Table 1. Participant demographics.

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] • Uses MyChart for tracking right dosing and frequency for medications • Manual pill box used in the morning and afternoon/evening	• Uses reminders on continuous glucose monitor and compression boot app devices • 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] • Uses MyChart for tracking right dosing and frequency for medications • 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] • Used to track medications on an app	• No current health experience tracking or reporting • 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] • 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 • 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 • Manual pill box used in the morning and afternoon/evening	• Uses health app for tracking glucose (<30 min/d) • 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

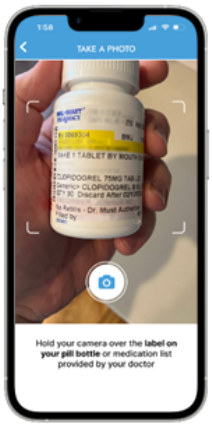
Other improvements to the Care4Today Connect app were discussed during the focus group. Areas of concern cited by participants included font sizes, number of fields, and amount of typing required to log into insurance portals to add medication data. One of the recommendations was a multiplatform/multidevice 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).

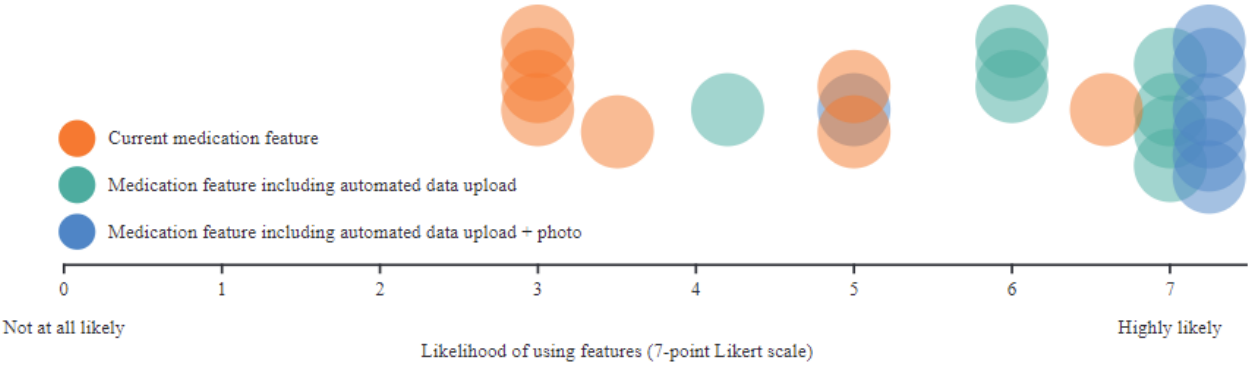
Adding Medication Data Feature

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.

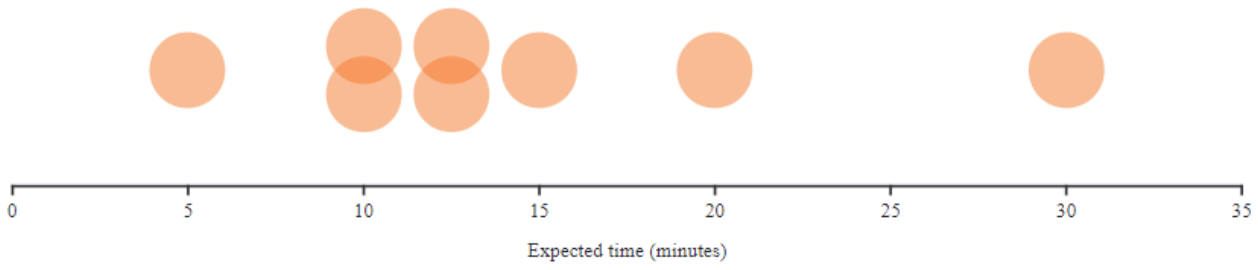
> *I want to start tracking my symptoms when and where they occur because my doctor does not believe me when I tell him.* [Female participant, 40-49, PAD]
>
> *I often forget what happened last week or last month, so I don't discuss my old symptoms with my doctor.* [Female participant, 60-69, CAD and 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.

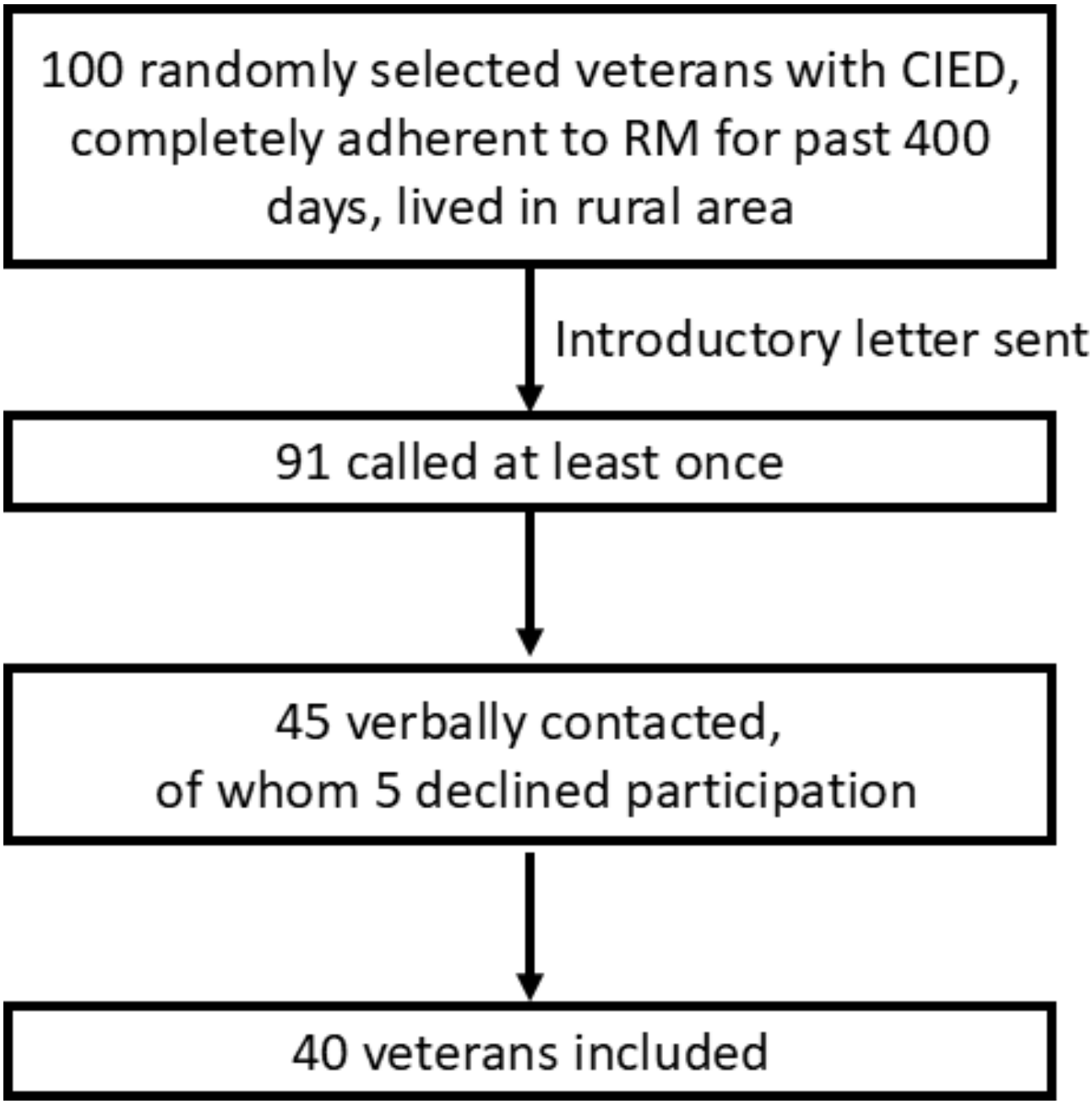


Table . Characteristics of veterans interviewed (n=40).

		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

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

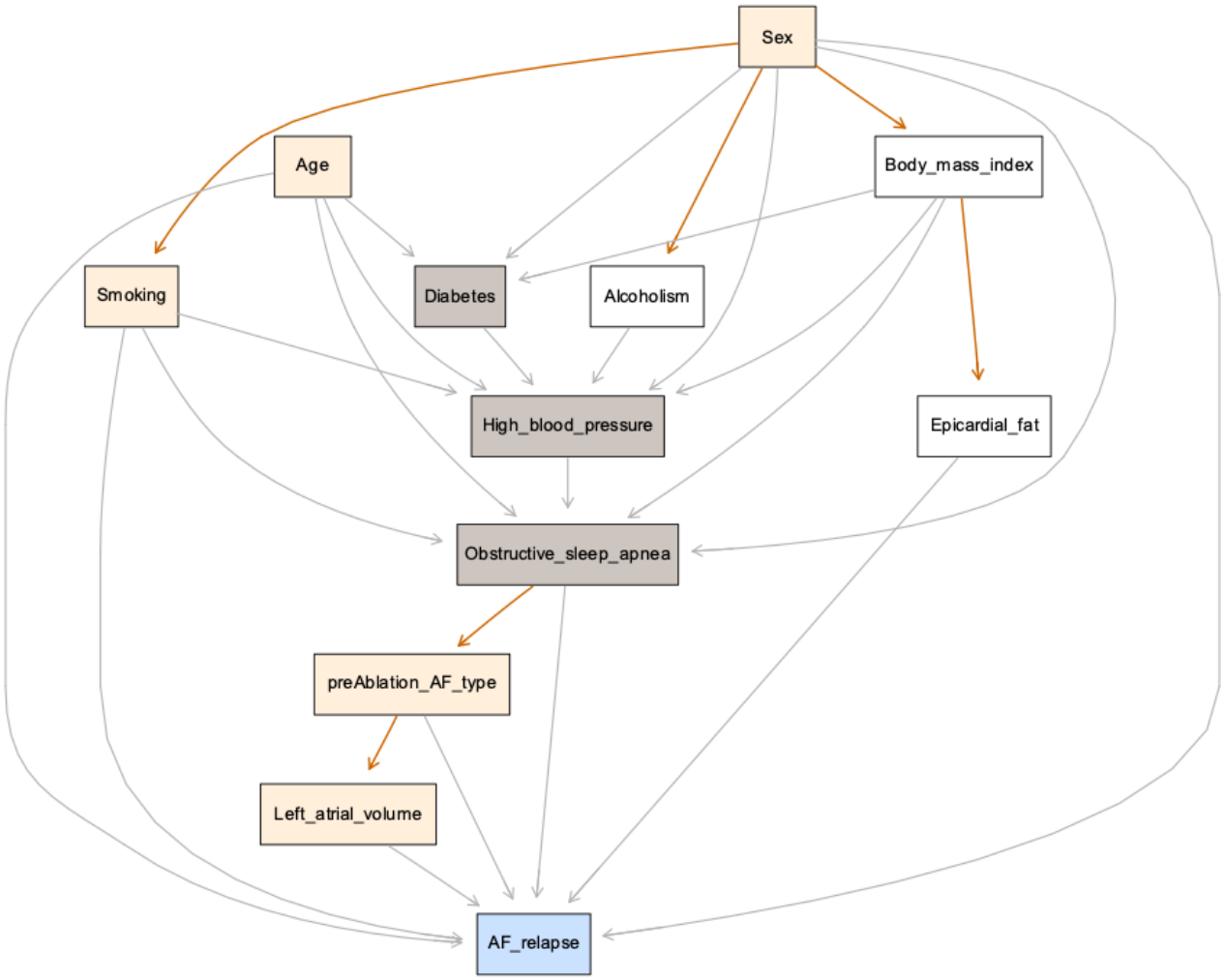
This study adheres to the ethical guidelines of the Declaration of Helsinki, including its later amendments. It has been approved by the Health Ethics Commission of the Western Lisbon Hospital Center, with the approval number 2117. All patients provided written informed consent before this study for 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.

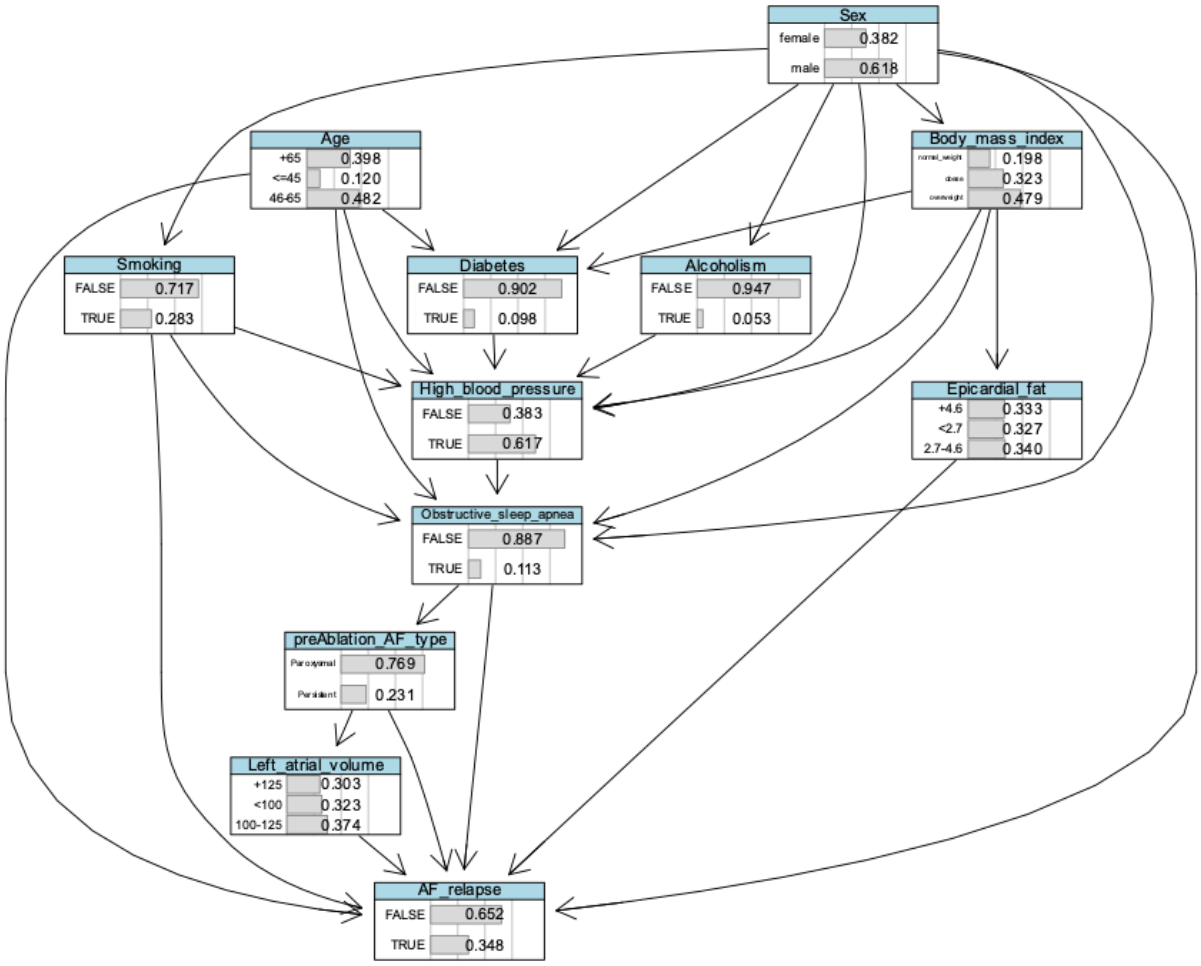
Figure 1. Bayesian network structure with nodes (boxes) representing the analyzed demographic and clinical variables. Grey nodes represent diseases with known associated risk factors, namely diabetes, high blood pressure, and obstructive sleep apnea. Beige nodes represent the 5 atrial fibrillation (AF) 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)

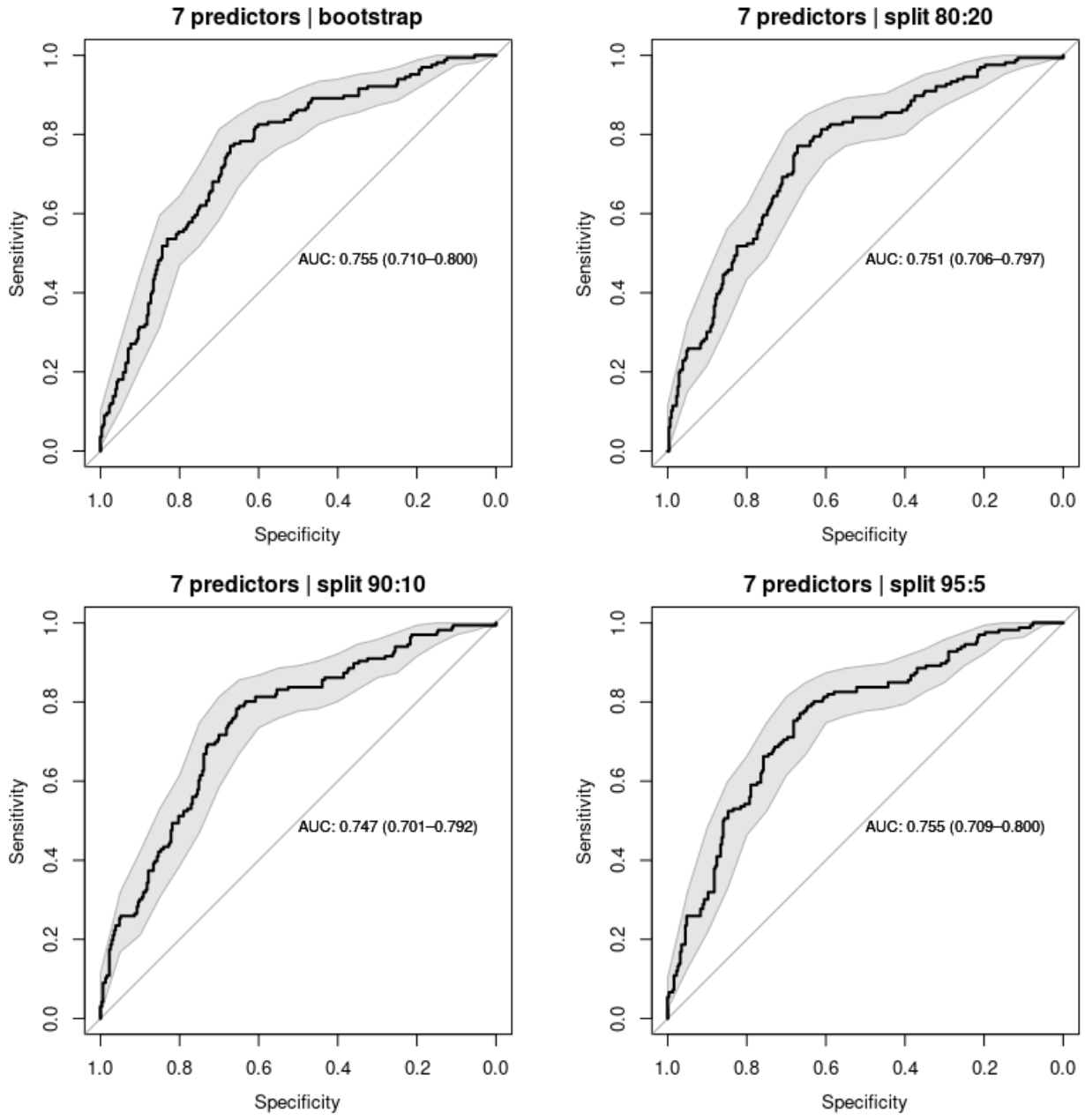
[a]Square brackets indicate that the end point is included in the range, and parentheses indicate that the end point is not included in the range.

[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.



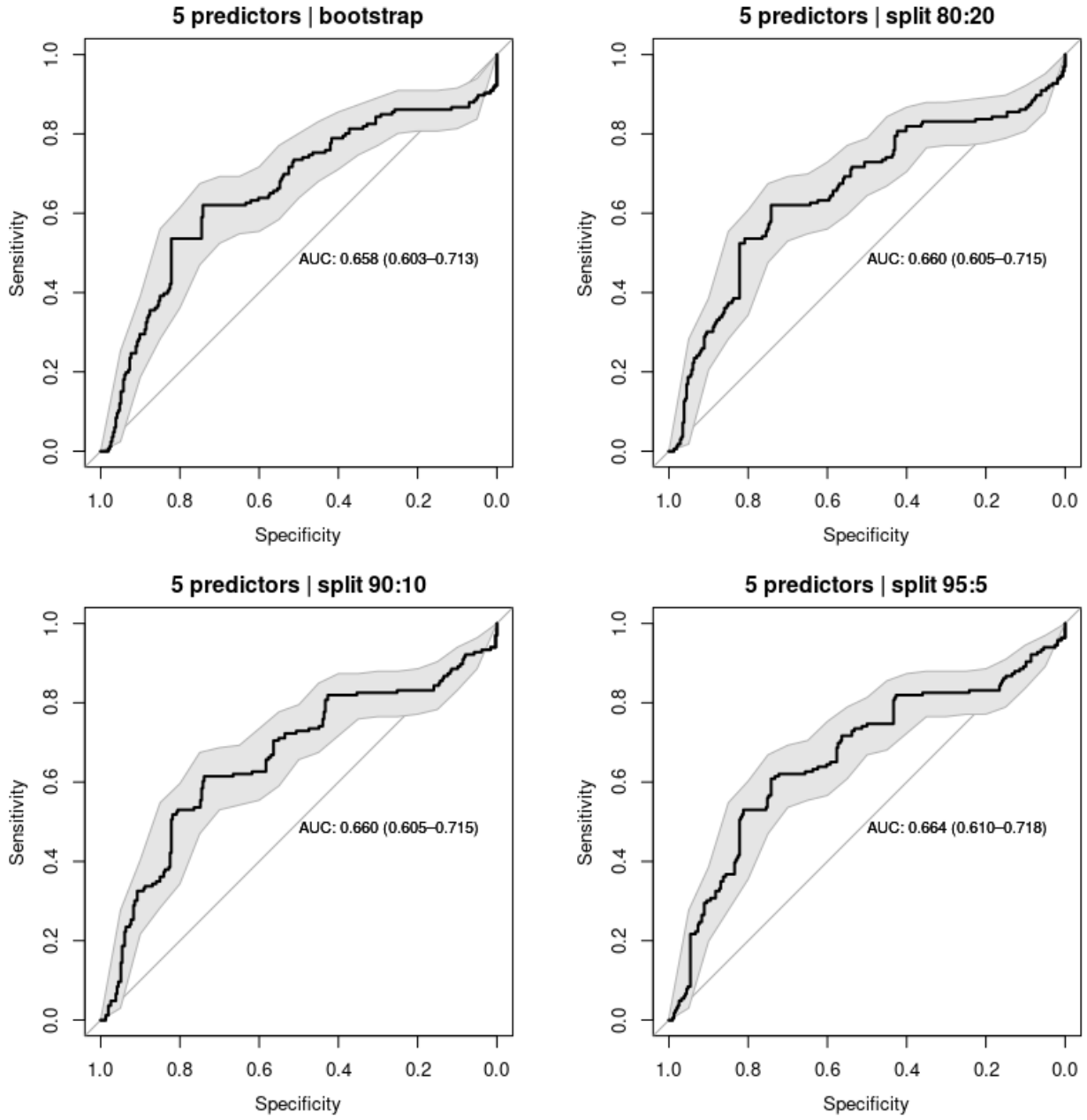
The 5 Predictors

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.

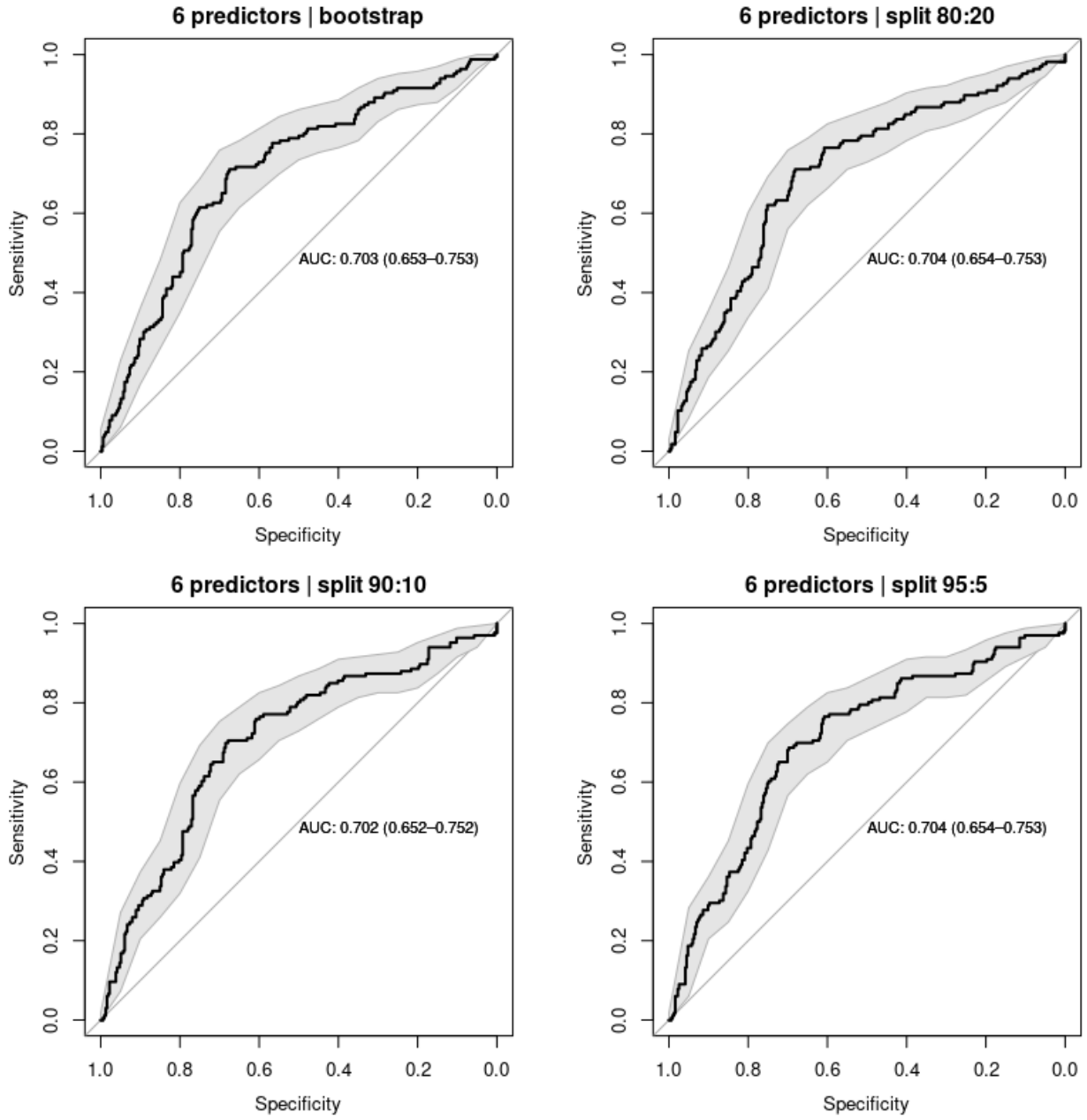


Table 3 presents a comparative analysis of the three models developed using 5, 6, and 7 predictors, respectively. As shown, the AUC-ROC progressively increases with the addition of predictors, indicating improved model performance. Furthermore, the 95% CI narrows as the number of predictors increases, 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

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org