JMIR Cardio

Electronic, mobile, digital health approaches in cardiology and for cardiovascular health
Volume 8 (2024) ISSN 2561-1011 Editor in Chief: Naomi Cahill, PhD, MSc, Scientific Editor at JMIR Publications, Canada

Contents

Original Papers

Review

Original Paper

Effects of a Web-based Weight Management Education Program on Various Factors for Overweight and Obese Women: Randomized Controlled Trial

Yunmin Han[1], MSc; Hoyong Sung[2], PhD; Geonhui Kim[3], PhD; Yeun Ryu[1], MSc; Jiyeon Yoon[4], MSc; Yeon Soo Kim[1], MD, PhD

[1]Department of Physical Education, Seoul National University, Seoul, Republic of Korea
[2]Department of Military Kinesiology, Korea Military Academy, Seoul, Republic of Korea
[3]Department of Aviation Sports, Korea Air Force Academy, Cheongju, Republic of Korea
[4]Department of Kinesiology, Iowa State University, Ames, IA, United States

Corresponding Author:
Yeon Soo Kim, MD, PhD
Department of Physical Education
Seoul National University
Seoul, Gawnak Gu, Gawankro 1
71-1, 408
Seoul, 08826
Republic of Korea
Phone: 82 2 880 7894
Email: kys0101@snu.ac.kr

Abstract

Background: Mediated diet and exercise methods yield effective short-term weight loss but are costly and hard to manage. However, web-based programs can serve many participants, offering ease of access and cost-efficiency.

Objective: This study aimed to compare the effectiveness of a web-based weight management program through web-based education alone (MINE) or combined with tailored video feedback (MINE Plus) with a control (CO) group.

Methods: This intervention included 60 Korean women with overweight and obesity (BMI≥23 kg/m^2) aged 19 years to 39 years old. We randomly allocated 60 participants to each of 3 groups: (1) MINE group (web-based education video and self-monitoring app), (2) MINE Plus group (web-based education video, self-monitoring app, and 1:1 tailored video feedback), and (3) CO group (only self-monitoring app). Web-based education included nutrition, physical activity, psychological factors, medical knowledge for weight loss, goal setting, and cognitive and behavioral strategies. Tailored feedback aimed to motivate and provide solutions via weekly 10-minute real-time video sessions. The intervention lasted 6 weeks, followed by a 6-week observation period to assess the education's lasting effects, with evaluations at baseline, 6 weeks, and 12 weeks. A generalized linear mixed model was used to evaluate time and group interactions.

Results: In the intention-to-treat analysis including all 60 participants, there were significant differences in weight change at 6 weeks in the MINE and MINE Plus groups, with mean weight changes of –0.74 (SD 1.96) kg (P=.03) and –1.87 (SD 1.8) kg (P<.001), respectively, while no significant change was observed in the CO group, who had a mean weight increase of 0.03 (SD 1.68) kg (P=.91). After 12 weeks, changes in body weight were –1.65 (SD 2.64) kg in the MINE group, –1.59 (SD 2.79) kg in the MINE Plus group, and 0.43 (SD 1.42) kg in the CO group. There was a significant difference between the MINE and MINE Plus groups (P<.001). Significant group × time effects were found for body weight in the MINE and CO groups (P<.001) and in the MINE Plus and CO groups (P<.001), comparing baseline and 12 weeks. Regarding physical activity and psychological factors, only body shape satisfaction and health self-efficacy were associated with improvements in the MINE and MINE Plus groups (P<.001).

Conclusions: This study found that the group receiving education and tailored feedback showed significant weight loss and improvements in several psychological factors, though there were differences in the sustainability of the effects.

Trial Registration: Korea Disease Control and Prevention Agency (KDCA) KCT0007780: https://cris.nih.go.kr/cris/search/detailSearch.do/22861

(JMIR Cardio 2024;8:e42402) doi:10.2196/42402

KEYWORDS

weight loss; obesity; health education; self-management; health promotion; tailored feedback; web-based intervention; behavior change

Introduction

Obesity is linked to a wide range of diseases and increases the risk of morbidity and mortality [1]. In South Korea, there has been a steady yearly increase in the obesity trend, from 29.7% in 2009 to 36.3% in 2019 [2]. The social and economic burdens of obesity on health care systems worldwide are significant, with costs associated with obesity-related diseases continuing to rise [3]. Addressing obesity effectively not only improves individual health outcomes but also reduces these broader economic impacts [4]. As a result, this trend has led to a growing medical burden, which has become a significant social issue [5]. It has become an ongoing public health concern due to its increasing prevalence each year [6]. Furthermore, it is associated with a high risk of metabolic syndrome and chronic diseases such as type 2 diabetes, high blood pressure, and cardiovascular disease [7,8]. Studies have shown that people with obesity have a lower quality of life than those who do not [9], causing mental health problems such as an increased risk of depression [10]. Therefore, maintaining a healthy weight is essential to prevent physical and psychological health problems.

Treatments for obesity encompass various strategies, including psychological management. Research has shown that behavior modification programs targeting weight loss not only lead to significant weight reduction but also a decrease in depression [11]. Furthermore, studies focusing on the psychological aspects of weight loss, such as depression, anxiety, and quality of life, reveal that women with overweight or obesity are more likely to view their body image pessimistically than men [12,13]. Addressing these psychological challenges through obesity education and management programs grounded in social cognitive theory has been proven effective [14].

Over the years, various weight loss programs, including behavioral therapies, have been carried out to address obesity [15,16]. Exercise and nutrition are usually addressed by experts face to face in most conventional weight loss studies. Recently, web-based research has become increasingly popular due to its ability to save time, costs, and human resources compared with previous face-to-face research. However, these studies have several limitations, such as participants needing more motivation, data collection problems, and low attrition rates [17]. The sustainability of health behaviors postintervention is critical for long-term weight management success. Exploring strategies to maintain and support these behaviors beyond the intervention period is essential, underscoring the importance of follow-up and continued engagement [18]. Moreover, digital transformation in the exercise and medical sectors has become increasingly inevitable due to advancements in artificial intelligence and digital health care. The digital era has accelerated the growth of the online home training and telemedicine market [19]. Therefore, more studies are needed to increase the sample size, include various target groups, and confirm continuous effects to overcome the limitations.

According to a meta-analysis of behavioral change programs, study duration has varied from 2 weeks to 78 weeks (mean 26 weeks). However, given that only 5 of the 35 digital-based studies included a follow-up period, the number of studies with such periods was very limited [20]. Furthermore, studies that did not include goal setting and feedback showed low research quality. Therefore, to emphasize the effectiveness of behavioral change programs, high-quality research with feedback and goal setting is needed [21].

The primary aim of this study was to compare the weight loss effects of web-based education and feedback among 3 groups (2 interventions and 1 control) over a given period and to determine whether participants could implement what they learned and achieve weight loss. The second aim was to compare whether the effects of physical activity and psychological factors persisted in these groups during the observation period.

Methods

Recruitment and Participants

Participants in this study were recruited from Seoul National University students and staff members through advertisements (eg, email, flyers, and social media). The advertisements included information on the purpose of the study, data collection methods, and benefits offered to participants. According to the Korean Society for the Study of Obesity, a BMI of 23 kg/m^2 to 24.9 kg/m^2 is considered overweight, a BMI of 25 kg/m^2 to 29.9 kg/m^2 is considered first-degree obesity, a BMI of 30 kg/m^2 to 34.9 kg/m^2 is considered second-degree obesity, and a BMI ≥35 kg/m^2 is defined as severe obesity [22]. Eligible participants were young women aged 19 years to 39 years, with a BMI >23 kg/m^2 according to Asian standards, who were able to listen to and write Korean, and who could use the Internet and smartphone devices.

Exclusion criteria were a loss of more than 10% of body weight in the past 6 months, previous obesity surgery, and pregnancy. People diagnosed with severe mental illness or cardiovascular metabolism or who were receiving medication that could affect weight loss were also excluded. All participants submitted their informed consent before enrollment.

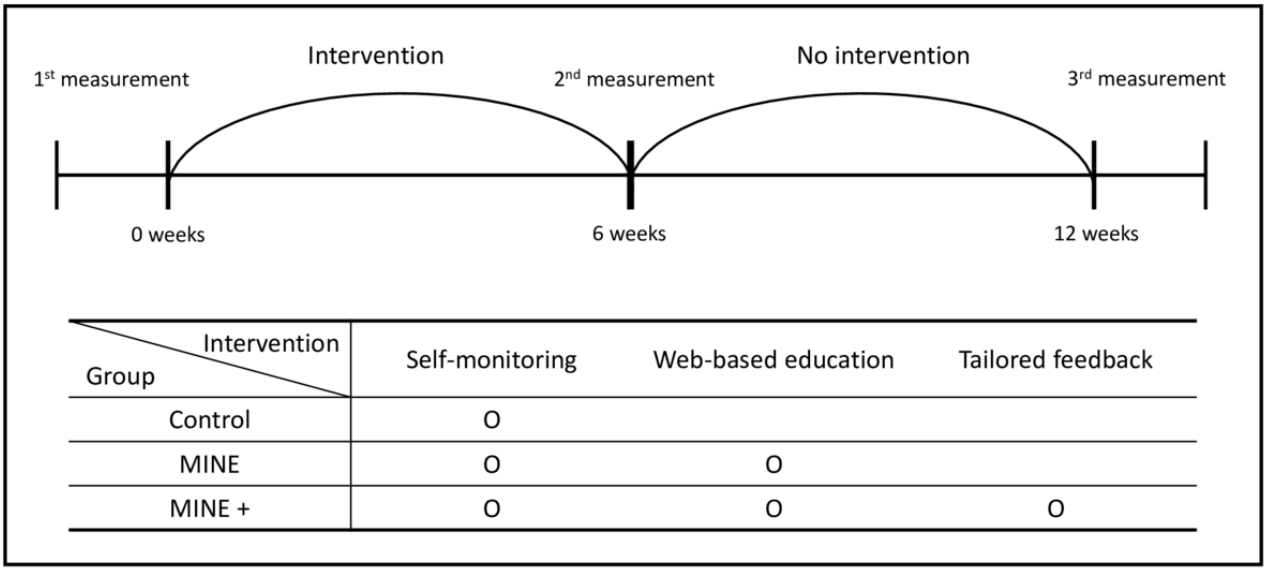
Interventions

The web-based program carried out in this study is referred to as “MINE,” which is a combination of “Mind,” “Medicine,” “Nutrition,” and “Exercise.” Participants received necessary education on weight loss in all 4 of these fields (Figure 1).

Figure 1. Program curriculum characteristics including research and methods related to nutrition, psychology, exercise, and medicine for weight management.

This program does not force physical activity on the participants nor interfere with their diet. Instead, through education and feedback as the intervention, it aims to change behavior to achieve weight loss by empowering the participants to make better choices on their own. The group that received only web-based education was called "MINE," the group that received web-based education and tailored feedback was called "MINE Plus," and the waiting list group was called "Control" (Figure 2).

Figure 2. Experimental period and group type.



Group \ Intervention	Self-monitoring	Web-based education	Tailored feedback
Control	O		
MINE	O	O	
MINE +	O	O	O

The sample size was calculated using GPower version 3.1. The significance level was .05, the effect size was 0.25, and the power was 0.8. The minimum number of participants was 42, the dropout rate was set at approximately 20% based on previous studies [23], and the number of participants was set at 60 and randomly assigned to 3 groups. We asked the participants to send their self-monitoring records once a week. However, there was no compulsion to send data even if the data were insufficient. The behavior change strategy was composed by incorporating the basic concepts of personalized cognitive behavioral therapy for obesity that were previously researched, as well as goal setting, social support, action planning, coping planning, and self-monitoring, which are frequently used in eHealth interventions. This information was integrated and utilized in conjunction with the education content and feedback [24,25]. The education content was divided into exercise, nutrition, psychology, and medical areas to establish knowledge and behavioral change strategies necessary for weight loss. Educational materials were produced by referring to previous scientifically proven studies and were verified by qualified experts in the field. The web-based education videos lasted 15 minutes to 20 minutes; details are shown in Table 1. In addition, these videos were delivered via video links and materials for each week through chat rooms organized by the groups.

Table 1. Educational content order and topics.

Number	Educational sessions	Field
1	Setting my diet	Nutrient
2	Setting my workout	Exercise
3	How to control my appetite	Psychology
4	Understanding fat burning system	Medicine
5	Healthy nutrients vs worst nutrients	Nutrient
6	Understanding the benefits of exercise	Exercise
7	Dealing with other people's perceptions and getting rid of stress	Psychology
8	Understanding the digestive system	Medicine
9	Practical diet recipes	Nutrient
10	How to exercise at home or the workplace	Exercise
11	How to deal with eating out and appointments	Psychology
12	Understanding the endocrine system and metabolism related to weight control	Medicine
13	Making your own sustainable diet plan	Nutrient
14	Making your own sustainable exercise goals	Exercise
15	Establishing long-term weight management strategies based on social cognitive theory	Psychology
16	Developing a self-check list and a plan for improving health indicators	Medicine

The feedback was based on the protocol of the web-based weight loss program called "POWeR" developed by Dennison et al [26] and a social cognitive theory strategy for obesity treatment developed by Dalle Grave et al [24]. Before proceeding with the feedback, basic lifestyle and weight loss experience information was collected through an online survey. The survey included information such as usual mealtimes, meal volume, sleep, health information, weight loss history, difficulties with losing weight, and personal goals. The feedback schedule was then implemented by entering the ID at the desired time through the online form and entering the Zoom video conference at the indicated time. Coaching sessions were conducted once a week, lasting approximately 15 minutes to 20 minutes each. Both groups (MINE and MINE Plus) received weekly educational materials and alarm messages about watching and practicing educational videos through group chat rooms. After receiving education, it was considered that the education was completed by submitting quizzes related to the content. In addition, a smart electronic weighing scale was provided as a research incentive. Of the 2 groups who did not receive feedback during the observation period, the MINE group received feedback, and the control group received educational videos and feedback after the end of the follow-up period.

Outcomes

The study lasted for a total of 12 weeks, consisting of a 6-week intervention period followed by a 6-week follow-up period. Participants' demographic data included age, marital status, and education level. The main outcome was to confirm the results related to body weight change at 6 weeks and 12 weeks after baseline. The secondary outcome was to confirm the level of physical activity, eating attitudes, satisfaction with body shape, and health self-efficacy.

The participants were examined in the laboratory after signing a consent form. Height, body weight, body composition, waist circumference, and blood pressure (BP) were measured. Height was measured using a BSM230 (Biospace), and weight and body fat were measured using an InBody 720 (Biospace). Participants were advised to avoid heavy meals, water, and intense physical activity for at least 2 hours before undergoing the InBody measurement to minimize potential measurement bias. Waist circumference was marked in centimeters using a tape measure from the middle of the lower rib to the upper iliac. For the BP measurement, an arm cuff for adults was placed around the left upper arm (Watch BP 03, Microlife AG). After measuring BP twice at rest, the average of the 2 values was calculated for the final BP index. We conducted 4 surveys after collecting the basic physical information. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) and calculated as metabolic equivalents (METs) [27]. In addition, the Eating Attitudes Test-26 (EAT-26), developed by Garner and Garfinkel [28], was used to measure attitudes toward eating, while the Body Shape Questionnaire (BSQ), developed by Cooper et al [29], was used to measure satisfaction with body shape. In addition, the Self-Rated Abilities for Health Practices (SRAHP) scale developed by Becker et al [30] was used as a measure of health self-efficacy [30]. Validity and reliability were verified for all questionnaires translated into Korean [31-33].

The EAT score was obtained by subtracting 65 points from the total score. A tendency for eating problems is assumed if the final score is 18 or higher, while a severe eating problem is

considered to exist if the final score is 25 or higher [33]. The BSQ is a 32-point to 192-point scale, with higher scores indicating a greater sense of obesity and lower overall satisfaction with body shape [29,34]. The SRAHP is a 24-point to 96-point scale, with higher scores indicating a greater sense of self-efficacy regarding one's health [30,31]. After the measurements were completed, participants were randomly allocated to the groups through a lottery. Precautions for participation in the study were delivered orally.

The same methods and questionnaires were used for every measurement. In addition, a program satisfaction survey on web-based education programs was conducted, and tailored feedback was provided on the intermediate measurement after the intervention. The survey included satisfaction with the overall program, satisfaction with personal feedback, and recommendations for the program (see Multimedia Appendix 1); it was modified by referring to the satisfaction survey in the preliminary study developed by Yip et al [35].

Statistical Analysis

Intention-to-treat analysis was performed for all outcomes at 6 weeks and 12 weeks. To determine differences in the variables, we set the time (baseline, 6 weeks, 12 weeks) as the repeated factors for body weight, body composition, blood pressure, physical activity, mental health score, and physical health score.

The Shapiro-Wilk normality test for each group was performed. Only the MINE Plus group did not show a normal distribution (P<.01). Therefore, for variables that did not have a normal distribution, such as body weight, BMI, waist circumference, and physical activity, a generalized linear mixed model was used to fit log-transformed data [36]. A linear mixed model method was used for the remaining variables satisfying the normality test [37]. These methods set fixed and random effects, and the interaction effect of time and group on the outcome was confirmed as outcomes of the variables. All statistical analyses were conducted using R studio (version 4.0.3). The mean and SD were calculated using descriptive statistical analysis, and the significance level was set at P<.05.

Ethical Considerations

The study received approval from the international review board at Seoul National University, Seoul, South Korea (SNU IRB NO. 2109/002-007). All participants provided written consent to participate, with a clear process established for medical referral and reporting any potential harm arising from their participation. The research also emphasized privacy and confidentiality protection for human subjects by ensuring that all study data were either anonymous or de-identified, accompanied by a brief description of the protective measures in place. To safeguard the privacy and confidentiality of participants, additional details on these protections were provided. As compensation for their involvement, participants received a smart scale valued at approximately US $20.

Results

Participant Characteristics

Of the 88 participants, 28 were excluded because their BMI did not exceed 23 kg/m^2 or they did not meet the inclusion criteria (Figure 3). A total of 60 people were assigned to the 3 groups, 20 per group, through random allocation; for their characteristics, see Table 2. All 60 participants were analyzed using the intention-to-treat method. In addition, missing values for unmeasured participants and participants who dropped out were analyzed as the last measured data through the last observation carried forward method [38].

Figure 3. Flow diagram of participants through the trial.

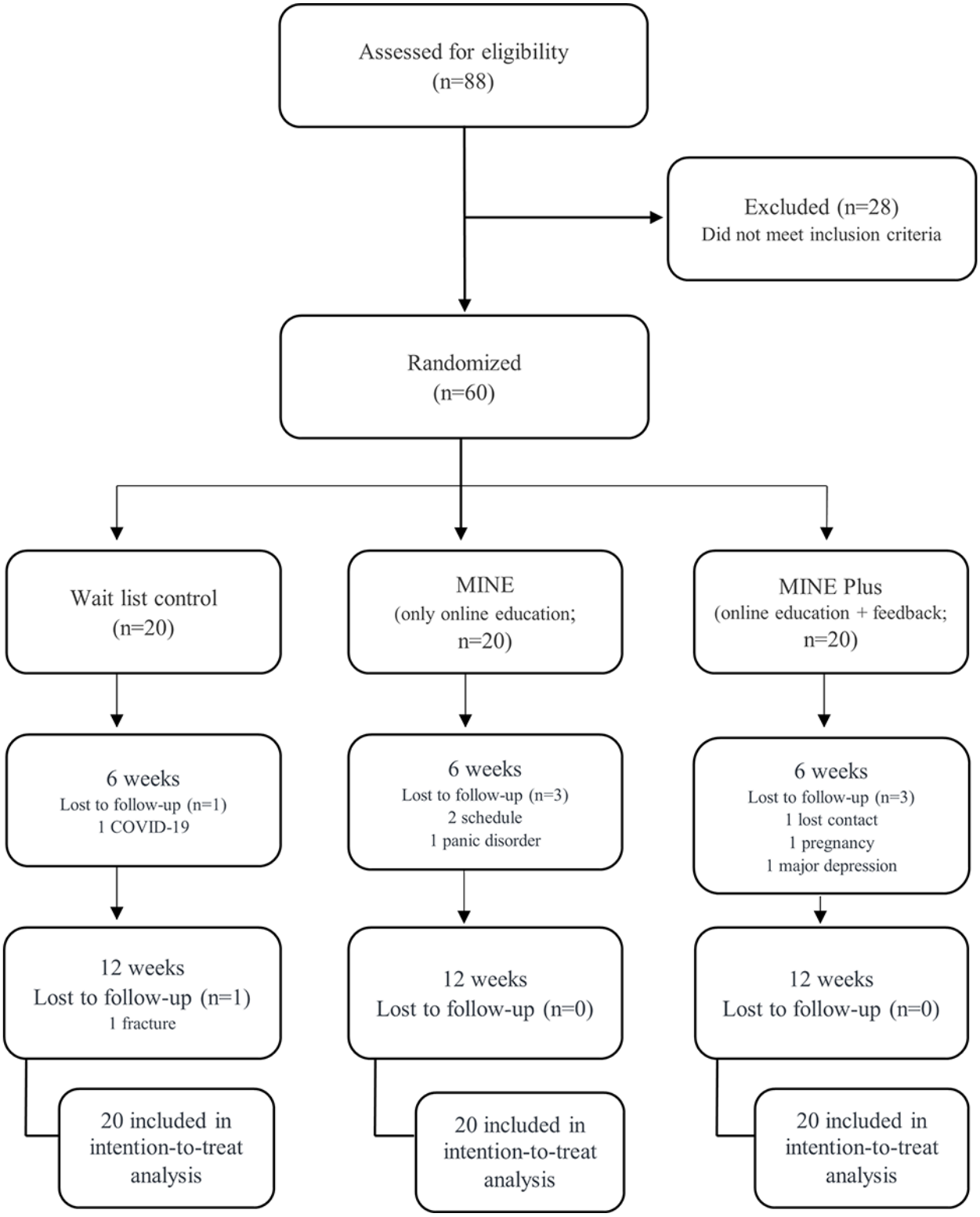


Table 2. Baseline characteristics of participants (N=60).

Characteristic	MINE[a] (n=20), mean (SD)	MINE Plus[b] (n=20), mean (SD)	Control (n=20), mean (SD)
Age (years)	29.7 (5.44)	28.2 (3.75)	30.3 (5.06)
Weight (kg)	65.82 (7.23)	70.64 (13.07)	66.2 (7.08)
BMI (kg/m^2)	25.3 (2.09)	26.5 (4.11)	25.2 (2.00)
Body fat percentage (%)	36.8 (4.59)	38.5 (5.98)	36.5 (4.68)
Waist circumference (cm)	84.0 (6.21)	86.9 (9.81)	84.6 (6.54)
Blood pressure (mm Hg)			
Systolic	117 (10.63)	114 (7.98)	117 (10.98)
Diastolic	74 (8.03)	72 (6.44)	76 (10.44)
Resting heart rate (bpm)	78 (13.22)	75 (10.31)	82 (11.29)
Physical activity (METs[c]/week)	1717 (1368)	1246 (1226)	1853 (2153)
EAT-26[d]	6.79 (16.84)	10.53 (12.64)	6.84 (17.89)
BSQ[e]	110.55 (32.96)	116.63 (26.48)	118.79 (27.56)
SRAHP[f]	70.88 (12.53)	74.02 (11.74)	72.88 (10.66)

[a]MINE: only online education.

[b]MINE Plus: online education + tailored feedback.

[c]MET: metabolic equivalent.

[d]EAT-26: Eating Attitudes Test-26 (Korean version).

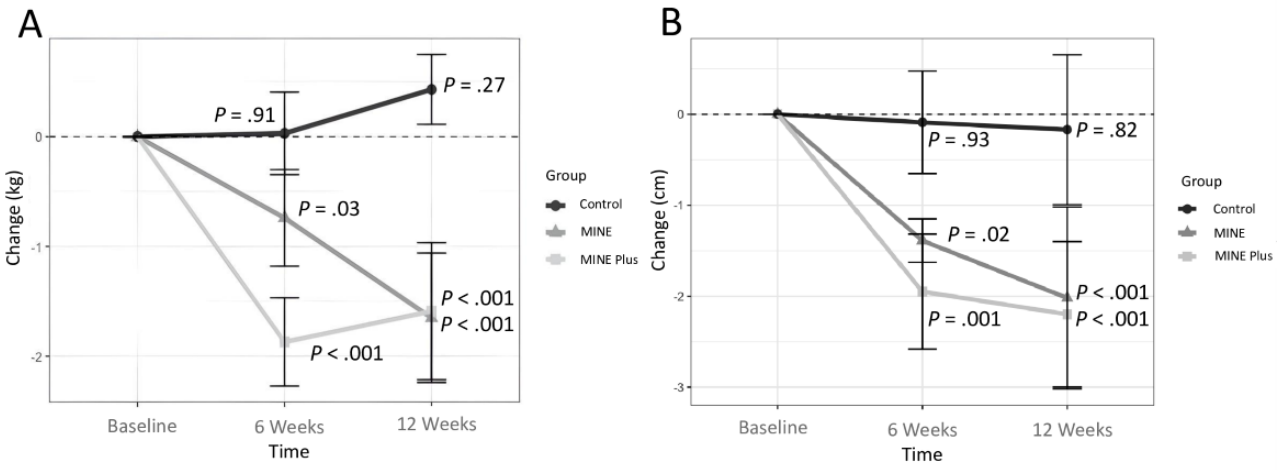
[e]BSQ: Body Shape Questionnaire (Korean version).

[f]SRAHP: Self-Rated Abilities of Health Practices (Korean version).

Weight Loss

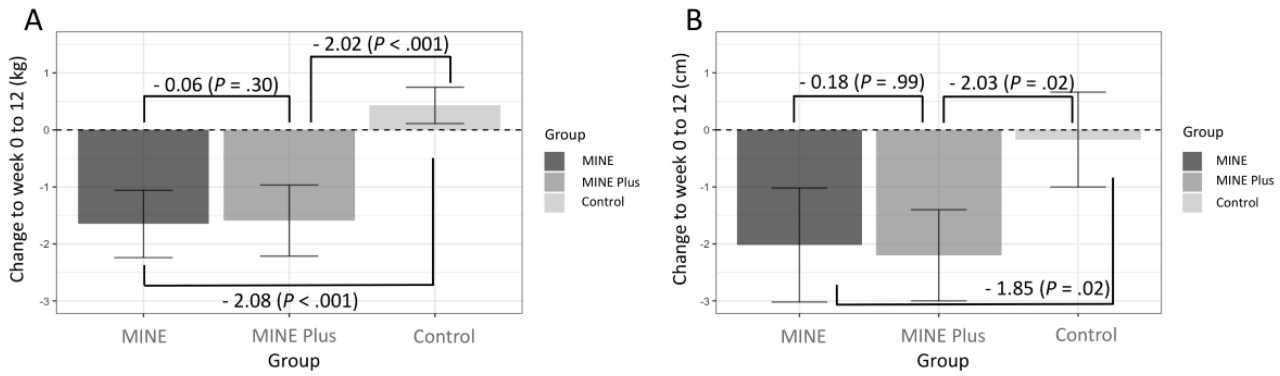
From baseline to week 6, the MINE group showed a significant mean weight reduction of –0.74 (SD 1.96) kg (*P*=.03), the MINE Plus group showed a significant mean weight reduction of –1.87 (SD 1.8) kg (*P*<.001), and the control group showed a mean weight increase of 0.03 (SD 1.68) kg (*P*=.91; Figure 4). From week 6 to week 12, the mean weight reduction was –0.91 (SD 2.2) kg (*P*<.001) for the MINE group, while the mean increases in weight were 0.28 (SD 1.72) kg (*P*=.24) for the MINE Plus group and 0.41 (SD 1.42) kg (*P*=.23) for the control group. Comparing the baseline with week 12, the MINE group showed a total mean reduction of –1.65 (SD 2.64) kg (*P*<.001), while the MINE Plus group showed a significant mean weight reduction of –1.59 (SD 2.79) kg (*P*<.001). In contrast, the control group showed a mean weight increase of 0.43 (SD 1.42) kg (*P*=.27). From baseline to Week 12, waist circumference significantly decreased by a mean –2.02 (SD 4.47) cm (*P*<.001) in the MINE group and –2.2 (SD 3.58) cm (*P*<.001) in the MINE Plus group, but the decrease of a mean –0.17 (SD 3.72) cm in the control group was not significant (*P*=.82).

Figure 4. Mean changes in (A) body weight and (B) waist circumference at baseline, 6 weeks, and 12 weeks, with *P* values for differences from baseline at 6 weeks and 12 weeks. MINE: only online education; MINE Plus: online education + tailored feedback.



The interaction between time and group for weight change was not significant from baseline to week 12 after the intervention in the MINE and MINE Plus groups (*P*=.30; Figure 5). The MINE and control groups did not show any interaction effect until week 6 of the intervention (*P*=.11), although there was a significant effect (*P*<.001) compared with week 12 (see Multimedia Appendix 2). The MINE Plus and control groups showed a significant interaction effect when comparing baseline with week 6 and baseline with week 12 (*P*<.001), but this was absent from week 6 to week 12 (*P*=.98).

Figure 5. Comparison of (A) body weight and (B) waist circumference interaction effects between time and groups, with the *P* values for the differences between groups from baseline to 12 weeks. MINE: only online education; MINE Plus: online education + tailored feedback.



Physical Activity Changes

Physical activity was calculated as the percentage of participants meeting the minimum amount of physical activity recommended by the World Health Organization (600 METs/week) [39]. No significant changes in physical activity levels were observed in any group during the postmeasurement and follow-up periods (Table 3).

Table 3. Differences in weight, waist circumference (WC), physical activity, and psychological factors between groups (N=60).

Characteristic	MINE[a] (n=20)	*P* value	MINE Plus[b] (n=20)	*P* value	Control (n=20)	*P* value
Weight at baseline (kg), mean (SD)	65.82 (7.23)	—[c]	70.64 (13.07)	—	66.2 (7.08)	—
Weight changes (kg), mean (SD)						
Baseline to postintervention	–0.74 (1.96)	.03	–1.87 (1.8)	<.001	0.03 (1.68)	.91
Postintervention to follow-up	–0.91 (2.2)	<.001	0.28 (1.72)	.24	0.41 (1.16)	.23
Baseline to follow-up	–1.65 (2.64)	<.001	–1.59 (2.79)	<.001	0.43 (1.42)	.27
WC at baseline (cm), mean (SD)	84.0 (6.21)	—	86.9 (9.81)	—	84.6 (6.54)	—
WC changes (cm), mean (SD)						
Baseline to postintervention	–1.39 (3.57)	.02	–1.95 (2.83)	<.001	–0.09 (2.52)	.93
Postintervention to follow-up	–0.63 (3.31)	.29	–0.25 (2.2)	.71	–0.08 (3.13)	.89
Baseline to follow-up	–2.02 (4.47)	<.001	–2.2 (3.58)	<.001	–0.17 (3.72)	.82
Physical activity at baseline (≥600 METs[d]), n (%)	16 (80)	—	13 (65)	—	14 (70)	—
Physical activity changes (%)[e], mean (SD)						
Baseline to postintervention	–15 (1.83)	.13	0 (2.29)	.60	–5 (1.97)	.60
Postintervention to follow-up	10 (2.24)	.31	0 (2.29)	.77	0 (2.29)	.97
Baseline to follow-up	5 (1.97)	.81	0 (1.62)	.62	–5 (1.97)	.58
EAT-26[f,g] at baseline	6.79 (16.84)	—	10.53 (12.64)	—	6.84 (17.89)	—
EAT-26 changes, mean (SD)						
Baseline to postintervention	4.95 (10.44)	.07	3.3 (10.95)	.22	–0.1 (9.64)	.97
Postintervention to follow-up	–0.3 (15.3)	.91	1.25 (8.88)	.64	0.5 (8.26)	.85
Baseline to follow-up	4.65 (18.73)	.09	4.55 (12.65)	.09	0.4 (9.68)	.88
BSQ[h,i] at baseline, mean (SD)	110.55 (32.96)	—	116.63 (26.48)	—	118.79 (27.56)	—
BSQ changes, mean (SD)						
Baseline to postintervention	–10.45 (29.03)	.02	–10.35 (20.07)	.03	0.7 (18.43)	.88
Postintervention to follow-up	–2.25 (13.64)	.62	0.3 (13.29)	.62	2.25 (13.88)	.62
Baseline to follow-up	–12.7 (33.67)	.006	–10.05 (16.99)	.03	2.95 (13.9)	.52
SRAHP[j,k] at baseline, mean (SD)	70.88 (12.53)	—	74.02 (11.74)	—	72.88 (10.66)	—
SRAHP changes, mean (SD)						
Baseline to postintervention	5.65 (9.83)	.003	7.2 (8.68)	<.001	1.8 (6.83)	.34
Postintervention to follow-up	1.1 (9.9)	.56	0.25 (6.54)	.89	2.05 (5.95)	.28
Baseline to follow-up	6.75 (12.34)	<.001	7.45 (7.49)	<.001	3.85 (5.63)	.04

[a]MINE: only online education.

[b]MINE Plus: online education + feedback.

[c]Not applicable.

[d]MET: metabolic equivalent.

[e]The minimum World Health Organization recommended amount of 600 METs minutes per week.

[f]EAT: Eating Attitudes Test.

[g]Higher scores indicate more negative eating attitudes.

[h]BSQ: Body Shape Questionnaire.

[i]Higher scores indicate lower satisfaction.

[j]SRAHP: Self-Rated Abilities of Health Practices.

[k]Higher scores indicate better self-efficacy.

Psychological Factor Changes

After 12 weeks, the EAT-26 scores increased by 4.65 (SD 18.73; *P*=.09) for the MINE group, 4.55 (SD 12.65; *P*=.09) for the MINE Plus group, and 0.4 (SD 9.68; *P*=.88) for the control group. After 12 weeks, the BSQ scores decreased by –12.7 (SD 33.67; *P*=.006) for the MINE group and –10.05 (SD 16.99; *P*=.03) for the MINE Plus group but increased by 3.85 (SD 5.63; *P*=.04) for the control group, which means the MINE and MINE Plus groups experienced greater satisfaction with their body shape. By 6 weeks, the BSQ scores for the MINE and MINE Plus groups had significantly decreased by –10.45 (SD 29.03; *P*=.02) and –10.35 (SD 20.07; *P*=.03), respectively. However, there was no significant decrease from 6 weeks to 12 weeks, the period during which there was no intervention. The SRAHP scores increased by 6.75 (SD 12.34; *P*<.001) for the MINE Plus group and 7.45 (SD 7.49; *P*<.001) for the control group after 12 weeks. After the intervention, the SRAHP scores for both the MINE and MINE Plus groups increased significantly, by 5.65 (SD 9.83; *P*=.003) and 7.2 (SD 8.68; *P*<.001), respectively, but did not increase significantly until 12 weeks after the intervention.

Discussion

Principal Findings

This study confirmed changes in weight loss and psychological factors for women with overweight and obesity aged in their 20s and 30s through a web-based, weight-related, behavioral change program. Both the MINE group, who received only web-based education, and the MINE Plus group, who received web-based education and feedback, showed significant weight loss. Regarding psychological factors, significant positive changes were observed only in body shape satisfaction and health management self-efficacy.

Of the 60 participants in this study, 54 completed the postintervention and follow-up measurements, resulting in a dropout rate of ~10%. Mobile- or web-based, health-related interventions are usually reported to have high dropout rates [25]. According to a systematic literature review by Kelders et al [40], the median duration of web-based health interventions was 10 weeks, and the mean adherence rate for web-based trial participants was 55%. However, the proportions of participants who received the education and submitted quizzes were 68% in the MINE group and 84% in the MINE Plus group (Multimedia Appendix 3). This showed higher compliance than that observed in previous studies, due to continuous motivation and the accessibility of educational content. Therefore, this study shows that even web-based experimental studies can have low dropout rates.

Comparison With Previous Work

Compared with previous studies, a study by Baer et al [23] showed an average weight loss of –3.1 (95% CI –3.7 to –2.5) kg over 12 months in an integrated intervention group with online programs and health managers. In addition, studies based on other online platforms and coaching programs showed an average weight loss of –1.57 (95% CI –1.92 to –1.22) kg over 24 weeks [15]. In this study, the average 12-week loss in the intervention groups was –1.65 (SD 2.64) kg in the MINE group and –1.59 (SD 2.79) kg in the MINE Plus group, which was a significant loss compared with the weight difference observed in previous studies.

In our study, weight loss and lifestyle changes were documented through app use records, feedback, and satisfaction surveys; however, objective indicators could not be confirmed. A lack of power, insufficient time to produce results, and insufficient sensitivity of measurement tools to detect minor differences could be reasons for no differences in physical activity or cardiovascular outcomes [21]. To increase physical activity, it is considered essential to identify individual vulnerabilities through “just-in-time” feedback and to set goals suitable for participants, taking into consideration the available time and circumstances [41]. It also requires specific guidance and clarification on when, where, and how to act [42]. Considering this situation, interventions and elements that can increase physical activity and make it habitual should be further developed.

There was no significant difference in weight reduction in the MINE Plus group in the absence of the intervention. The group who received feedback showed a higher rate of weight loss until the intervention period, but it is likely that independence decreased after the intervention ended due to the absence of coaching. Participants may experience various problems such as anxiety and decreased self-confidence if they receive feedback for a certain period and then abruptly discontinue it for a short period [43,44]. In contrast, it seemed that the MINE group showed a continuous form of weight loss by managing goal setting based on the educational content. Therefore, it is necessary to establish a more effective feedback methodology when mediating feedback and proceed to provide feedback until the participants are independent and weight management is achieved through long-term feedback. However, considering that the waist circumference of the MINE Plus group decreased more than that of the MINE group over the entire period, it is inferred that more effective health management was conducted.

A previous study confirmed the positive effect of body shape satisfaction using Internet-based mindfulness interventions [45]. In this study also, there was a significant improvement in body shape satisfaction as the education program provided strategies to the participants on how to cope with social stigma or stress, similar to mindfulness strategies. Finally, it seems that continuous motivation and encouragement were helpful through tailored feedback.

In a study confirming the relationship between weight loss and health self-efficacy in young adults, self-efficacy increased through weight management education and showed effective results in promoting weight loss [46]. In this study, participants managed their weight with methods that could set and achieve goals through education. It showed that providing goals through feedback, complimenting them, and encouraging them were helpful. However, no other positive results could be confirmed after week 6. Previous research has indicated that the effects of tailored feedback are controversial, with effectiveness diminishing over extended periods of observation, showing no significant difference from groups who did not receive

intervention [47,48]. The provision of goals is significantly correlated with self-efficacy, which leads to the creation of sustainable self-set goals [49]. However, it appears that, after the intervention and education period, the participants' self-efficacy and self-set goal development did not expand without being given goals. Therefore, there is a need for interventions in which participants can continue to be motivated and achieve goals.

Psychological factors for weight loss identified for young women in this study were to treat obesity when young and relieve psychological pressure through healthy weight loss. Obesity treatment should proceed in a healthy way that prevents access to compulsive and distorted knowledge and can be self-applied [50]. Therefore, the program was organized in a way to empower the participants to improve their habits and perceptions through education rather than by forced intervention. This study progressed with education and feedback based on various theories, including the social cognitive theory. Although it is difficult to determine which theories and methods worked best in the program, the approach of integrated theory will be more critical in health education [21]. Furthermore, future research should investigate and compare the effectiveness of various delivery methods in online programs to identify which theories and strategies produce the most substantial outcomes for participants or patients [51]. Furthermore, collective and institutional efforts should be accompanied at the national level to improve long-term weight management and dietary and activity environments [52].

Strengths and Limitations

This study has several strengths. First, the program was constructed by approaching the treatment of obesity in a multifactorial manner rather than by considering just one factor. Tools such as nutrition, physical activity, psychological and medical knowledge education, and tailored feedback were used to improve lifestyle, and the program was designed so that participants could improve themselves. Second, this approach can reduce time, expenses, and human resources. Since this program is only conducted online, participants can proceed with web-based education and feedback according to their preferred place and time. Furthermore, educational videos can be reused later. Feedback also allows participants and moderators to participate at any time and place of their choice. As they turned on the camera and communicated in real time, the participants could see each other’s facial expressions, understand emotions, and build bonds. Third, existing behavioral change programs were used to supplement and construct obesity treatment strategy elements in the program. Various behavioral change factors, such as goal setting, self-monitoring, and self-control, which are essential and proven effective, were continuously added to the educational content and feedback to maximize effectiveness. However, there are several limitations to this study. First, there was no long-term follow-up of the participants. Obesity is an area that requires continuous management, and it is necessary to set a period for long-term follow-up. According to the meta-analysis by Beleigoli et al [17], web-based digital interventions have shown effectiveness in the short term but not in the long term. Therefore, for long-term management, it seems necessary to combine offline management with digital interventions. Second, the use of various objective indicators was insufficient. The indicators for physical activity and psychological factors in this study were in the form of questionnaires. Therefore, more objective data should be collected through the observation of physical activity using accelerometers or various blood and biomarker data. Third, the nutritional information collected through self-monitoring was not evaluated due to measurement bias. It appears that future research will require accurate collection and analysis of nutritional information. Fourth, only university undergraduates, graduate students, and faculty members participated in this study. From a demographic perspective, most of them were highly educated participants. Therefore, the data from this program cannot be generalized to all other young adult women in the real world. Last, there was no significant change in physical activity because of the spread of COVID-19 in South Korea during the experimental period.

Conclusions

This study demonstrated the efficacy of a web-based education program, with and without tailored feedback, in promoting weight loss and enhancing psychological well-being through self-managed diet and exercise modifications. The 2 groups who received the intervention experienced significant weight loss over time, yet the magnitude of weight reduction varied across periods. Although improvements were observed in various psychological factors, these psychological improvements did not persist in the absence of the intervention. This indicates a need to integrate social support, incentives, or motivation in future digital health interventions and underscores the importance of interdisciplinary research in this field.

Acknowledgments

We extend our gratitude to all participants who took part in the experiment. The authors attest that there was no use of generative artificial intelligence (AI) technology in the generation of text, figures, or other informational content of this manuscript.

Authors' Contributions

YH designed the research, managed the investigation process, developed the methodology, and contributed to writing the manuscript. HS was responsible for analyzing the data. JY and GK were involved in data collection and participated in editing the manuscript. YR contributed to editing the manuscript. YSK provided supervision throughout the research process and contributed to reviewing the manuscript. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Web-based education program and tailored feedback satisfaction.
[DOCX File , 29 KB - cardio_v8i1e42402_app1.docx]

Multimedia Appendix 2
Supplementary table (Effects of interaction between time and groups).
[DOCX File , 44 KB - cardio_v8i1e42402_app2.docx]

Multimedia Appendix 3
Quiz submission rate.
[DOC File , 33 KB - cardio_v8i1e42402_app3.doc]

Multimedia Appendix 4
CONSORT-EHEALTH checklist.
[PDF File (Adobe PDF File), 1376 KB - cardio_v8i1e42402_app4.pdf]

References

1. Lenz M, Richter T, Mühlhauser I. The morbidity and mortality associated with overweight and obesity in adulthood: a systematic review. Dtsch Arztebl Int 2009 Oct;106(40):641-648 [FREE Full text] [doi: 10.3238/arztebl.2009.0641] [Medline: 19890430]
2. Yang YS, Han B, Han K, Jung J, Son JW, Taskforce Team of the Obesity Fact Sheet of the Korean Society for the Study of Obesity. Obesity Fact Sheet in Korea, 2021: Trends in obesity prevalence and obesity-related comorbidity incidence stratified by age from 2009 to 2019. J Obes Metab Syndr 2022 Jun 30;31(2):169-177 [FREE Full text] [doi: 10.7570/jomes22024] [Medline: 35770450]
3. Nagi MA, Ahmed H, Rezq MAA, Sangroongruangsri S, Chaikledkaew U, Almalki Z, et al. Economic costs of obesity: a systematic review. Int J Obes (Lond) 2024 Jan 26;48(1):33-43. [doi: 10.1038/s41366-023-01398-y] [Medline: 37884664]
4. Cleo G, Glasziou P, Beller E, Isenring E, Thomas R. Habit-based interventions for weight loss maintenance in adults with overweight and obesity: a randomized controlled trial. Int J Obes (Lond) 2019 Feb 23;43(2):374-383. [doi: 10.1038/s41366-018-0067-4] [Medline: 29686382]
5. Song HJ, Hwang J, Pi S, Ahn S, Heo Y, Park S, et al. The impact of obesity and overweight on medical expenditures and disease incidence in Korea from 2002 to 2013. PLoS One 2018;13(5):e0197057 [FREE Full text] [doi: 10.1371/journal.pone.0197057] [Medline: 29746509]
6. Liu B, Du Y, Wu Y, Snetselaar LG, Wallace RB, Bao W. Trends in obesity and adiposity measures by race or ethnicity among adults in the United States 2011-18: population based study. BMJ 2021 Mar 16;372:n365 [FREE Full text] [doi: 10.1136/bmj.n365] [Medline: 33727242]
7. Abdelaal M, le Roux CW, Docherty NG. Morbidity and mortality associated with obesity. Ann Transl Med 2017 Apr;5(7):-161 [FREE Full text] [doi: 10.21037/atm.2017.03.107] [Medline: 28480197]
8. Aronow WS. Association of obesity with hypertension. Ann Transl Med 2017 Sep;5(17):-350 [FREE Full text] [doi: 10.21037/atm.2017.06.69] [Medline: 28936444]
9. Ul-Haq Z, Mackay DF, Fenwick E, Pell JP. Meta-analysis of the association between body mass index and health-related quality of life among adults, assessed by the SF-36. Obesity (Silver Spring) 2013 Mar 16;21(3):E322-E327 [FREE Full text] [doi: 10.1002/oby.20107] [Medline: 23592685]
10. Luppino FS, de Wit LM, Bouvy PF, Stijnen T, Cuijpers P, Penninx BWJH, et al. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. Arch Gen Psychiatry 2010 Mar;67(3):220-229. [doi: 10.1001/archgenpsychiatry.2010.2] [Medline: 20194822]
11. Linde JA, Simon GE, Ludman EJ, Ichikawa LE, Operskalski BH, Arterburn D, et al. A randomized controlled trial of behavioral weight loss treatment versus combined weight loss/depression treatment among women with comorbid obesity and depression. Ann Behav Med 2011 Feb 28;41(1):119-130 [FREE Full text] [doi: 10.1007/s12160-010-9232-2] [Medline: 20878292]
12. Pereira-Miranda E, Costa PRF, Queiroz VAO, Pereira-Santos M, Santana MLP. Overweight and obesity associated with higher depression prevalence in adults: a systematic review and meta-analysis. J Am Coll Nutr 2017;36(3):223-233. [doi: 10.1080/07315724.2016.1261053] [Medline: 28394727]

13. Vancini RL, Rayes ABR, Lira CABD, Sarro KJ, Andrade MS. Pilates and aerobic training improve levels of depression, anxiety and quality of life in overweight and obese individuals. Arq. Neuro-Psiquiatr 2017 Dec;75(12):850-857. [doi: 10.1590/0004-282x20170149]
14. Bennion KA, Tate D, Muñoz-Christian K, Phelan S. Impact of an Internet-based lifestyle intervention on behavioral and psychosocial factors during postpartum weight loss. Obesity (Silver Spring) 2020 Oct 04;28(10):1860-1867 [FREE Full text] [doi: 10.1002/oby.22921] [Medline: 32888250]
15. Beleigoli A, Andrade AQ, Diniz MDF, Ribeiro AL. Personalized web-based weight loss behavior change program with and without dietitian online coaching for adults with overweight and obesity: randomized controlled trial. J Med Internet Res 2020 Nov 05;22(11):e17494 [FREE Full text] [doi: 10.2196/17494] [Medline: 33151151]
16. Johnston BC, Kanters S, Bandayrel K, Wu P, Naji F, Siemieniuk RA, et al. Comparison of weight loss among named diet programs in overweight and obese adults: a meta-analysis. JAMA 2014 Sep 03;312(9):923-933. [doi: 10.1001/jama.2014.10397] [Medline: 25182101]
17. Beleigoli AM, Andrade AQ, Cançado AG, Paulo MN, Diniz MDFH, Ribeiro AL. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. J Med Internet Res 2019 Jan 08;21(1):e298 [FREE Full text] [doi: 10.2196/jmir.9609] [Medline: 30622090]
18. Walugembe DR, Sibbald S, Le Ber MJ, Kothari A. Sustainability of public health interventions: where are the gaps? Health Res Policy Syst 2019 Jan 15;17(1):8 [FREE Full text] [doi: 10.1186/s12961-018-0405-y] [Medline: 30646911]
19. Batsis JA, Dokko R, Naslund JA, Zagaria AB, Kotz D, Bartels SJ, et al. Opportunities to improve a mobile obesity wellness intervention for rural older adults with obesity. J Community Health 2020 Feb 05;45(1):194-200 [FREE Full text] [doi: 10.1007/s10900-019-00720-y] [Medline: 31486958]
20. Carraça E, Encantado J, Battista F, Beaulieu K, Blundell J, Busetto L, et al. Effective behavior change techniques to promote physical activity in adults with overweight or obesity: A systematic review and meta-analysis. Obes Rev 2021 Jul;22 Suppl 4(Suppl 4):e13258 [FREE Full text] [doi: 10.1111/obr.13258] [Medline: 33949778]
21. Jahangiry L, Farhangi MA. Obesity paradigm and web-based weight loss programs: an updated systematic review and meta-analysis of randomized controlled trials. J Health Popul Nutr 2021 Apr 08;40(1):16 [FREE Full text] [doi: 10.1186/s41043-021-00240-3] [Medline: 33832547]
22. WHO expert consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. The Lancet 2004 Jan;363(9403):P157-P163. [doi: 10.1016/s0140-6736(03)15268-3]
23. Baer HJ, Rozenblum R, De La Cruz BA, Orav EJ, Wien M, Nolido NV, et al. Effect of an online weight management program integrated with population health management on weight change: a randomized clinical trial. JAMA 2020 Nov 03;324(17):1737-1746 [FREE Full text] [doi: 10.1001/jama.2020.18977] [Medline: 33141209]
24. Dalle Grave R, Sartirana M, Calugi S. Personalized cognitive-behavioural therapy for obesity (CBT-OB): theory, strategies and procedures. Biopsychosoc Med 2020;14:5 [FREE Full text] [doi: 10.1186/s13030-020-00177-9] [Medline: 32175002]
25. Schroé H, Van Dyck D, De Paepe A, Poppe L, Loh WW, Verloigne M, et al. Which behaviour change techniques are effective to promote physical activity and reduce sedentary behaviour in adults: a factorial randomized trial of an e- and m-health intervention. Int J Behav Nutr Phys Act 2020 Oct 07;17(1):127 [FREE Full text] [doi: 10.1186/s12966-020-01001-x] [Medline: 33028335]
26. Dennison L, Morrison L, Lloyd S, Phillips D, Stuart B, Williams S, et al. Does brief telephone support improve engagement with a web-based weight management intervention? Randomized controlled trial. J Med Internet Res 2014 Mar 28;16(3):e95 [FREE Full text] [doi: 10.2196/jmir.3199] [Medline: 24681761]
27. Cleland CL, Hunter RF, Kee F, Cupples ME, Sallis JF, Tully MA. Validity of the global physical activity questionnaire (GPAQ) in assessing levels and change in moderate-vigorous physical activity and sedentary behaviour. BMC Public Health 2014 Dec 10;14(1):1255 [FREE Full text] [doi: 10.1186/1471-2458-14-1255] [Medline: 25492375]
28. Garner DM, Garfinkel PE. Socio-cultural factors in the development of anorexia nervosa. Psychol Med 1980 Nov 09;10(4):647-656. [doi: 10.1017/s0033291700054945] [Medline: 7208724]
29. Cooper PJ, Taylor MJ, Cooper Z, Fairbum CG. The development and validation of the body shape questionnaire. International Journal of Eating Disorders 1987 Jul;6(4):485-494. [doi: 10.1002/1098-108x(198707)6:4<485::aid-eat2260060405>3.0.co;2-o]
30. Becker H, Stuifbergen A, Oh HS, Hall S. Self-rated abilities for health practices: a health self-efficacy measure. Health Values: The Journal of Health Behavior, Education & Promotion 1993;17(5):42-50 [FREE Full text]
31. Lee J, Hong H, Park J. Validation study of the Korean Self Rated Abilities for Health Practices: health self-efficacy measure (K-SRAHP). Stress 2018 Sep 30;26(3):140-148. [doi: 10.17547/kjsr.2018.26.3.140]
32. Noh Y, Kim BW. The validation study of the body shape questionnaire (BSQ) in female university students. Korea Journal of Counseling 2005;6(4):1163-1174 [FREE Full text]
33. Rhee MK. A validation of the Korean version of Eating Attitude Test-26. Korean Journal of Psychosomatic Medicine 1998;9(2):153-163 [FREE Full text]
34. Lee S, Oh KJ. Validation study of the Sociocultural Attitudes towards Appearance Questionnaire in Korea. The Korean Journal of Clinical Psychology 2003;22(4):913-926 [FREE Full text] [doi: 10.1037/t08749-000]

35. Yip MP, Chang AM, Chan J, MacKenzie AE. Development of the Telemedicine Satisfaction Questionnaire to evaluate patient satisfaction with telemedicine: a preliminary study. J Telemed Telecare 2003 Jun 24;9(1):46-50. [doi: 10.1258/135763303321159693] [Medline: 12641893]
36. Lo S, Andrews S. To transform or not to transform: using generalized linear mixed models to analyse reaction time data. Front Psychol 2015 Aug;6:1171 [FREE Full text] [doi: 10.3389/fpsyg.2015.01171] [Medline: 26300841]
37. Hutchesson MJ, Collins CE, Morgan PJ, Callister R. An 8-week web-based weight loss challenge with celebrity endorsement and enhanced social support: observational study. J Med Internet Res 2013 Jul 04;15(7):e129 [FREE Full text] [doi: 10.2196/jmir.2540] [Medline: 23827796]
38. Elobeid MA, Padilla MA, McVie T, Thomas O, Brock DW, Musser B, et al. Missing data in randomized clinical trials for weight loss: scope of the problem, state of the field, and performance of statistical methods. PLoS One 2009 Aug 13;4(8):e6624 [FREE Full text] [doi: 10.1371/journal.pone.0006624] [Medline: 19675667]
39. 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 Dec 25;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
40. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
41. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res 2018 Mar 22;20(3):e106 [FREE Full text] [doi: 10.2196/jmir.8701] [Medline: 29567638]
42. Gollwitzer PM. Implementation intentions: Strong effects of simple plans. American Psychologist 1999 Jul;54(7):493-503. [doi: 10.1037/0003-066X.54.7.493]
43. Park J, Shea CH, Wright DL. Reduced-frequency concurrent and terminal feedback: a test of the guidance hypothesis. J Mot Behav 2000 Sep;32(3):287-296. [doi: 10.1080/00222890009601379] [Medline: 10975276]
44. Wulf G, Höß M, Prinz W. Instructions for motor learning: differential effects of internal versus external focus of attention. J Mot Behav 1998 Jun;30(2):169-179. [doi: 10.1080/00222899809601334] [Medline: 20037032]
45. Welsh EM, Sherwood NE, VanWormer JJ, Hotop AM, Jeffery RW. Is frequent self-weighing associated with poorer body satisfaction? Findings from a phone-based weight loss trial. J Nutr Educ Behav 2009 Nov;41(6):425-428 [FREE Full text] [doi: 10.1016/j.jneb.2009.04.006] [Medline: 19879499]
46. Roach JB, Yadrick M, Johnson JT, Boudreaux L, Forsythe WA, Billon W. Using self-efficacy to predict weight loss among young adults. J Am Diet Assoc 2003 Oct;103(10):1357-1359. [doi: 10.1016/s0002-8223(03)01072-1] [Medline: 14520258]
47. Ryan K, Dockray S, Linehan C. A systematic review of tailored eHealth interventions for weight loss. Digit Health 2019;5:2055207619826685 [FREE Full text] [doi: 10.1177/2055207619826685] [Medline: 30783535]
48. Sherrington A, Newham JJ, Bell R, Adamson A, McColl E, Araujo-Soares V. Systematic review and meta-analysis of internet-delivered interventions providing personalized feedback for weight loss in overweight and obese adults. Obes Rev 2016 Jun 07;17(6):541-551 [FREE Full text] [doi: 10.1111/obr.12396] [Medline: 26948257]
49. Phillips JM, Gully SM. Role of goal orientation, ability, need for achievement, and locus of control in the self-efficacy and goal--setting process. Journal of Applied Psychology 1997;82(5):792-802. [doi: 10.1037//0021-9010.82.5.792]
50. Dreber H, Reynisdottir S, Angelin B, Hemmingsson E. Who is the treatment-seeking young adult with severe obesity: a comprehensive characterization with emphasis on mental health. PLoS One 2015 Dec 22;10(12):e0145273 [FREE Full text] [doi: 10.1371/journal.pone.0145273] [Medline: 26694031]
51. Kozak AT, Buscemi J, Hawkins MAW, Wang ML, Breland JY, Ross KM, et al. Technology-based interventions for weight management: current randomized controlled trial evidence and future directions. J Behav Med 2017 Feb;40(1):99-111 [FREE Full text] [doi: 10.1007/s10865-016-9805-z] [Medline: 27783259]
52. Wadden TA, Tronieri JS, Butryn ML. Lifestyle modification approaches for the treatment of obesity in adults. Am Psychol 2020 Feb;75(2):235-251 [FREE Full text] [doi: 10.1037/amp0000517] [Medline: 32052997]

Abbreviations

BP: blood pressure
BSQ: Body Shape Questionnaire
EAT: Eating Attitudes Test
GPAQ: Global Physical Activity Questionnaire
MET: metabolic equivalent
SRAHP: Self-rated Abilities for Health Practices scale

Edited by A Mavragani; submitted 11.10.22; peer-reviewed by K Haas, J Antoun; comments to author 06.02.23; revised version received 22.05.23; accepted 23.03.24; published 18.04.24.

Please cite as:
Han Y, Sung H, Kim G, Ryu Y, Yoon J, Kim YS
Effects of a Web-based Weight Management Education Program on Various Factors for Overweight and Obese Women: Randomized Controlled Trial
JMIR Cardio 2024;8:e42402
URL: https://cardio.jmir.org/2024/1/e42402
doi:10.2196/42402
PMID:38635975

Original Paper

Cloud-Based Machine Learning Platform to Predict Clinical Outcomes at Home for Patients With Cardiovascular Conditions Discharged From Hospital: Clinical Trial

Phillip C Yang[1], MD; Alokkumar Jha[1], PhD; William Xu[2], BSc; Zitao Song[3], MSc; Patrick Jamp[4], BSc; Jeffrey J Teuteberg[1], MD

[1]Stanford University School of Medicine, Palo Alto, CA, United States

[2]Emory University, Atlanta, GA, United States

[3]North Carolina State University, Raleigh, NC, United States

[4]Electrical Engineering, University of California, Los Angeles, Mountain View, CA, United States

Corresponding Author:
Phillip C Yang, MD
Stanford University School of Medicine
300 Pasteur Dr # H2157 Stanford
Palo Alto, CA, 94305-2200
United States
Phone: 1 6508048828
Email: phillip@stanford.edu

Abstract

Background: Hospitalizations account for almost one-third of the US $4.1 trillion health care cost in the United States. A substantial portion of these hospitalizations are attributed to readmissions, which led to the establishment of the Hospital Readmissions Reduction Program (HRRP) in 2012. The HRRP reduces payments to hospitals with excess readmissions. In 2018, >US $700 million was withheld; this is expected to exceed US $1 billion by 2022. More importantly, there is nothing more physically and emotionally taxing for readmitted patients and demoralizing for hospital physicians, nurses, and administrators. Given this high uncertainty of proper home recovery, intelligent monitoring is needed to predict the outcome of discharged patients to reduce readmissions. Physical activity (PA) is one of the major determinants for overall clinical outcomes in diabetes, hypertension, hyperlipidemia, heart failure, cancer, and mental health issues. These are the exact comorbidities that increase readmission rates, underlining the importance of PA in assessing the recovery of patients by quantitative measurement beyond the questionnaire and survey methods.

Objective: This study aims to develop a remote, low-cost, and cloud-based machine learning (ML) platform to enable the precision health monitoring of PA, which may fundamentally alter the delivery of home health care. To validate this technology, we conducted a clinical trial to test the ability of our platform to predict clinical outcomes in discharged patients.

Methods: Our platform consists of a wearable device, which includes an accelerometer and a Bluetooth sensor, and an iPhone connected to our cloud-based ML interface to analyze PA remotely and predict clinical outcomes. This system was deployed at a skilled nursing facility where we collected >17,000 person-day data points over 2 years, generating a solid training database. We used these data to train our extreme gradient boosting (XGBoost)–based ML environment to conduct a clinical trial, *Activity Assessment of Patients Discharged from Hospital-I*, to test the hypothesis that a comprehensive profile of PA would predict clinical outcome. We developed an advanced data-driven analytic platform that predicts the clinical outcome based on accurate measurements of PA. Artificial intelligence or an ML algorithm was used to analyze the data to predict short-term health outcome.

Results: We enrolled 52 patients discharged from Stanford Hospital. Our data demonstrated a robust predictive system to forecast health outcome in the enrolled patients based on their PA data. We achieved precise prediction of the patients' clinical outcomes with a sensitivity of 87%, a specificity of 79%, and an accuracy of 85%.

Conclusions: To date, there are no reliable clinical data, using a wearable device, regarding monitoring discharged patients to predict their recovery. We conducted a clinical trial to assess outcome data rigorously to be used reliably for remote home care by patients, health care professionals, and caretakers.

(JMIR Cardio 2024;8:e45130) doi:10.2196/45130

KEYWORDS

smart sensor; wearable technology; moving average; physical activity; artificial intelligence; AI

Introduction

Background

Why are some discharged patients readmitted whereas others are not? Although often routine and uncomplicated, this transition of care is complex and, if not arranged properly, can lead to life-threatening consequences. Clearly, factors such as disease severity and the intensity of postdischarge care affect the risk of readmission; however, many other issues may also have substantial contributions [1]. The top contributory factors are (1) admission diagnosis: heart failure is the top cause of readmission, whereas other conditions, including sepsis, pneumonia, chronic obstructive pulmonary disease, and cardiac arrhythmia, are considered high risk; (2) insurance: Medicare and Medicaid patients have the highest rates of readmission; (3) patient demographics: race, sex, age, and income play a key role (eg, women who experience heart attacks and populations with lower-income status); and (4) patient engagement: patients who lack knowledge, skills, and confidence to manage their care have nearly double the average readmission rate [2,3].

The hospital readmission rate is approximately 20% in the United States, and the rates increase proportionately among those who are aged ≥50 years [4]. Our health care infrastructure is overburdened. Therefore, it is incumbent upon health care providers to develop risk stratification algorithms expeditiously to help predict which patients are at the highest risk for readmission. However, this determination is extremely difficult to make, particularly because the majority of the discharged patients will not become critically ill, require readmission, or die. Therefore, as we try to mitigate the risks of readmission, we need to do more than just predict the risk of readmission; we need to also tailor our home monitoring strategies to this risk [1]. Furthermore, this monitoring must not overwhelm health care providers or patients; rather, the aim should be to deliver smart, robust, and intelligent monitoring of patients convalescing at home.

Physical activity (PA) is one of the major determinants for overall clinical outcomes in chronic diseases, including diabetes, hypertension, hyperlipidemia, heart failure, and cancer, as well as mental health issues [5-8]. Moreover, these same comorbidities increase the risk of readmission. In 2017, the Centers for Disease Control and Prevention advocated adding PA as the fourth vital sign after heart rate (HR), blood pressure, and body temperature [9]. These developments underline the critical importance of PA in assessing the recovery of patients and, more importantly, indicate a clear need to measure PA quantitatively beyond the current questionnaire and survey methods [4,9,10]. PA is defined simply as any bodily movement produced by the skeletal muscles that result in energy expenditure [11]. However, it has been difficult historically to directly measure PA. It requires a dedicated laboratory to measure and perform a kinematic analysis. In addition, the measurement period is short and hard to monitor over time. Wearable technology and wireless data transmission have overcome these limitations and facilitate an accessible and long-term assessment of PA. A triaxial movement sensor was found to be a reliable, valid, and stable measurement of walking and daily PA in patients with chronic obstructive pulmonary disease [7]. Furthermore, a portable system for PA assessment in a home environment has been proposed [5]. These innovative systems provide novel and comprehensive real-time data for the evaluation of the health and quality of life of participants with limited mobility and chronic diseases. Finally, an estimate of step counts and energy expenditure strongly correlated with observed step counts and measured energy expenditure, using hip- and wrist-based Fitbit devices [6].

Development of an Advanced Data-Driven Analytic Platform

We developed an advanced data-driven analytic platform that predicts clinical outcomes based on accurate measurements of PA [10]. Artificial intelligence (AI) or machine learning (ML) analyzes the data to predict short- and long-term health outcomes. Although there is an overabundance of wearable devices (WDs) in the market, there are no known clinical outcome data that could be used reliably for home care by patients, health care professionals, or caretakers. In conjunction with AiCare Corp in San Jose, California, United States, we developed a platform consisting of the following key components: (1) a WD synced to an iPhone or app, (2) a web-based open application programming interface (API), (3) an AI and ML interface, and (4) a Health Insurance Portability and Accountability Act (HIPAA)–compliant Amazon Web Services (AWS) server environment. This platform was deployed at a skilled nursing facility where we collected >17,000 person-day data points (408,000 person-hour data points). These data provided the training set for our extreme gradient boosting (XGBoost) AI algorithm to correlate PA data to health outcomes in the *Activity Assessment of Patients Discharged from Hospital-I (ACT-I) clinical trial*. In this ACT-I trial, we enrolled 52 patients discharged from Stanford Hospital. Our data demonstrated a robust predictive system to forecast health outcomes in the enrolled patients based on their PA data. The clinical study generated a receiver operating characteristic (ROC) analysis with a sensitivity of 87% and a specificity of 79% in predicting the clinically significant events that were reported by the patients. Our comprehensive AI profiling of the PA of the discharged patients predicted their recovery or clinical deterioration to enable the precision guidance of appropriate and timely intervention during the 4-week follow-up period.

After considering various functionalities and requirements, the WD offered the most practical and compliant design solution to monitor discharged patients intelligently. However, the wearable technology for discharged patients should embody different applications and designs specific to their needs. In this paper, we will demonstrate these specifications in more detail. We will present AiCare's comprehensive technology solution consisting of a WD, Bluetooth low energy (BLE)–enabled iOS infrastructure, an ML algorithm to implement AI in patient care,

and API-enabled web technology to measure the daily activities of patients. In this study, remote data collection, robust XGBoost AI analysis, and the reliable prediction of clinical outcomes are reported.

Methods

Patient Recruitment

We screened patients discharged from Stanford Hospital general cardiology and advanced heart failure program.

Ethical Considerations

We obtained approval from the Stanford University Institutional Review Board (53805) and recruited patients discharged from Stanford Hospital. They were invited to participate in a research study to demonstrate the safety and feasibility of the AiCare platform. Informed consent was obtained from each participant who consented to primary data collection and secondary data analysis without additional consent. The privacy and confidentiality of participants are protected by a deidentified code that is assigned to each patient. No compensation was offered to participants.

ML Predictive Platform

Our comprehensive ML profile of the discharged patients was designed to predict their proper recovery to enable the precision guidance of timely intervention during the 4-week follow-up period. We developed an advanced data-driven XGBoost analytics platform to predict clinical outcomes based on accurate measurements of PA [10].

We compared several techniques to analyze our training data set, including logistic regression, naïve Bayes, support vector machine, and XGBoost. We measured precision, recall, F_1-score, area under the ROC curve (AUC-ROC), and average critical activity level, using different data sets. Throughout the analyses, XGBoost provided the highest area under the curve (AUC) values and other measurements.

We chose XGBoost because of its interpretability through the model training process, resistance to trivial features, and the reduced risk of overfitting. For our health care use case, model transparency was an important evaluation criterion. XGBoost visualized the feature prioritization and automatic weight assignments, which allowed us to explain the model insights to the stakeholders for solution adoption. Occasionally, there are noises in sensor data. To reduce the risk of overfitting, we experimented with max_depth of 2, 3, 4, and 5 and min_child_weight of 1, 2, 3, 4, and 5. On the basis of our list of PA-related input feature set and data volume (>17,000 person-day data points), the hyperparameters we used were max_depth of 3, learning_rate of 0.01, min_child_weight of 4, and n_estimators of 100. This achieved the balance of model accuracy, reducing the risk of overfitting and reaching a tolerable learning speed. We experimented with both XGBoost 1.4.1 and XGBoost 1.7.5. A mean absolute error of 1.7.x was introduced, which boosted the training algorithm convergence process. Some assumptions we made regarding the XGBoost model that should be considered include that the encoded integer values for each input variable have an ordinal relationship; it should not be assumed that all values are present. Our algorithm could handle missing values by default. In our tree-based algorithm, missing values were learned during the training phase.

Our AI platform predicted clinical outcome risk during the 4-week follow-up period using the continuous PA data stream. The PA features were measured by the number of occurrences of the multiples of g-force (1 g, 2 g, and 3 g) in each 1-hour time window. One hour was further divided into 7200 time intervals of 500 milliseconds each. Within each 500-millisecond period, the AiCare platform detected whether the minimum level of acceleration (1 g) had occurred. If yes, it increased the 1-g value by 1 count. Therefore, on an hourly basis, a restless user could potentially accumulate up to 7200 values of 1 g. The same detection and computational logic applied to 2-g and 3-g values. The directionless g-force was an aggregation of the g-forces in 3 axes (directionless g-force = $\sqrt{}$ [g-force_x^2 + g-force_y^2 + g-force_z^2]). This trial used the initial 72-hour period to build nonrisk baseline data and generated an alert when any deviation occurred, which indicated worsening health condition. The platform was able to detect the precursors of rehospitalization.

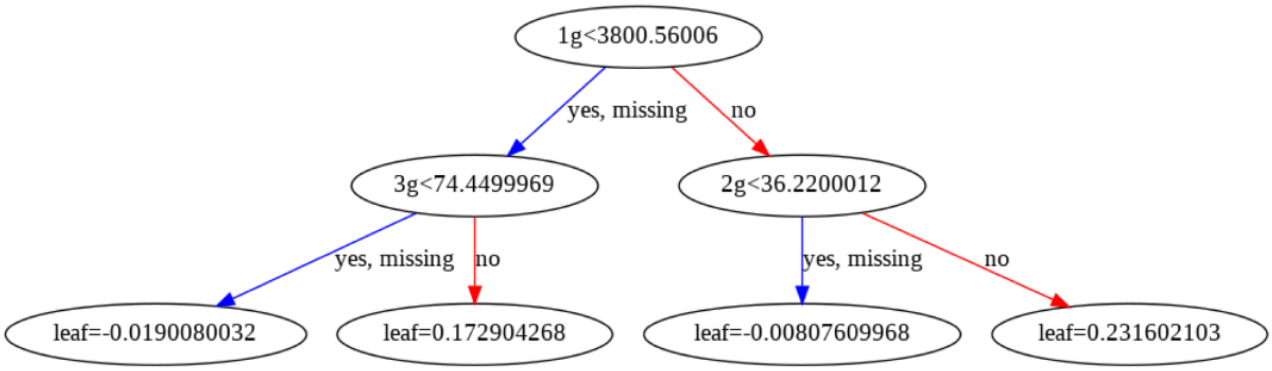
A decision tree ensemble–based multiclass classification approach was used to predict no risk, mild risk, and risk. A maximum tree depth of 3 levels was deployed. The intrinsic graph of the decision tree facilitated the explainability of the model. Figure 1 presents a sample decision tree from our model.

This decision tree visualization provides insight into the gradient boosting process. Figure 1 illustrates the importance and data coverage of each input feature (1 g, 2 g, and 3 g) and the decision-making process. We chose cross-entropy–based softprob objective (the loss function in the first term of the training objective equation presented after this paragraph) to predict the probabilities of 3 categories in the risk profile. Because of the tendency of the decision tree to bisect the data space and to overfit the training data when classes are not well separated, we introduced a regularization term to balance the bias-variance trade-off (the second term of the training objective equation presented after this paragraph).

(1)

To reduce false-positive results, we further enhanced the platform with the clinician's cognitive decision-based alerts. The AI model was trained continuously with the patients' up-to-date PA data. Besides the AI model (Figure 1), the AiCare prediction platform was designed for higher scalability. The inbound data pipeline supports open APIs that are hardware agnostic and integrate with the PA features collected from different hardware devices. The outbound patients' predictive risk profiles were streamed to various clinical applications to support broad clinical use cases. The ACT-I trial showed early signs of clinical efficacy with this low-cost noninvasive approach, enabling further scalability.

Figure 1. Sample decision tree.



Wireless Protocol

Considerations regarding the requirements of data collection, long-term use, power consumption, wireless transmission distance, legal radio frequency, home use, popularity, and cost led to choosing BLE as the optimal protocol for home care indoor use. The iPhone was connected to the cloud server via the standard wireless or cellular protocol.

ID System and Data Collection

The personalized data were anonymized using an internally specified ID system for data collection. A BLE media access control address for each band enabled this functionality. The patients wore the WD (a battery-powered smart band [Rockband; AiCare Corp]) at all times to enable continuous data collection (the Rockband has a battery life of 45 days for continuous use).

AiCare Technology

The AiCare platform enabled data collection and real-time analysis as described herein. The technology consisted of the following: (1) the low-cost and water-resistant Rockband with a battery life of 45 days for continuous use, (2) iPhone connectivity, (3) a cloud-enabled HIPAA-compliant AWS server, (4) open API architecture, (5) an ML and XGBoost interface, (6) an iPhone app, and (7) an AI-enabled COVID-19–specific questionnaire. This comprehensive platform was deployed in an iOS environment to analyze the patients' PA data. We assessed PA by a triaxial accelerometer, which provides the optimal solution between technological complexity and reliable measurement of PA. This service was designed to ensure a smart, safe, and secure environment enabled with real-time, intelligent, and timely tracking, detection, and analysis to promote a healthy and independent lifestyle for discharged patients.

Data Collection and Analysis

We used the Rockband for data collection. First, we defined the moving average (MA) of PA. The visualization of time series data obtained from AiCare's platform allowed us to (1) identify changes in energy level (EL) and movement percentage, (2) establish a personalized baseline for each discharged patient, and (3) understand the data trend to predict any deviation in daily activity pattern.

The MA of PA

The MA method is widely used to smooth out time series data by calculating the average values for a chosen period [4,12]. In this study, we used simple MA (SMA) to avoid the noisy measurements of the EL and movement percentage feature. Each data point was calculated using SMA in time series data and weighted equally. There was no need to set any weighting parameters such as the weighted or exponential MA method to generate SMA EL or movement percentages parameter. The SMA formula was defined as follows:

(2)

where P_k represented the data point at time k, and n was the chosen number of data points. A longer-term SMA was less sensitive in reflecting the change in data movement compared with a shorter-term SMA, which was used to highlight the major trends in time series data. A shorter-term MA was relatively faster to react to changes in trend, which was beneficial to applications that required a QR code. The adjustment of the value n in the equation measured the different effects of trend analysis. The specific features and the definitions of PA are outlined in Textbox 1.

Textbox 1. Studied features and definitions for physical activity patterns.

- Daytime energy level (EL): the EL obtained during the daytime period
- Nighttime EL: the EL obtained during the nighttime period
- Daily EL difference (ELD): the difference between daytime EL and nighttime EL
- Normalized ELD: the daily ELD in percentage values
- Daytime active percentage (AP): 100% minus daytime resting percentage (RP)
- Daytime RP: the percentage of zero movements during the daytime period
- Nighttime AP: 100% minus nighttime RP
- Nighttime RP: the percentage of zero movements during the nighttime period
- Daily active percentage difference: the difference between daytime AP and nighttime AP

Kinetic EL

In this study, the estimation of kinetic energy was used to describe the EL of PA. The original formula of kinetic energy is as follows:

$$\text{Kinetic energy} = \frac{1}{2} \times \text{mass} \times \text{velocity}^2 = \frac{1}{2} \times \text{mass} \times (\text{acceleration} \times \Delta t)^2 \quad (3)$$

The unit of kinetic energy is the joule (1 joule=1 kg m^2/s^2). To obtain a directionless measurement of acceleration from the accelerometer embedded in the Rockband, signal vector magnitude (SVM) was applied to calculate the overall magnitude of acceleration [13]:

(4)

where a_x, a_y, and a_z are the acceleration values from the triaxial accelerometer. In this system, the sampling rate of the accelerometer (Δt) is fixed and is equal to 50 Hz. The formula of kinetic energy can be rewritten as follows:

$$\text{Kinetic energy} = (\frac{1}{2} \times \text{mass} \times \Delta t^2) \times \text{acceleration}^2 = \text{Constant} \times SVM^2 \quad (5)$$

For each individual, the constant portion of this equation would be the same at any given time. Therefore, kinetic energy could be defined as SVM^2 ($m^2/s^2/kg$). As a result, the estimation of total EL from time 0 to time n is defined as follows:

$$\text{Total energy level} = SVM2(t1) + SVM2(t2) + \ldots + SVM2(tn) \quad (6)$$

After obtaining the estimated wake-up time and resting time from the cloud platform, we can calculate the total energy expenditure during the daytime period and nighttime period, respectively.

The daytime period is equal to the time between wake-up and resting times on the same day, and the nighttime period is equal to the time between resting time and subsequent wake-up time on the following day. We used daytime EL to estimate the total intensity of all PAs that happened during the daytime period and nighttime EL to represent the total intensity of all PAs that happened during the resting period. In addition, the daily EL difference (ELD) has been used to evaluate the daily PA changes:

$$\text{Daily energy level difference} = ELDaytime - ELNighttime \quad (7)$$

A positive value of daily ELD indicates that daytime EL is greater than nighttime EL. It may represent that an individual is active during the day or inactive (sleeps well) during the night, which is a healthy PA pattern. A negative value of daily ELD can be obtained when nighttime EL is greater than daytime EL. High nighttime EL may represent disrupted sleep patterns, and thus movements can be detected by the Rockband at night. A negative EL difference also means that the observed individual is inactive during the day. To compare the change in individual ELD, normalization has to be performed to convert the absolute values of ELD into the percentage of ELD, which is defined by the following equations:

(8)

$$\text{Active percentage (\%)} = 100\% - \text{resting percentage (\%)} \quad (9)$$

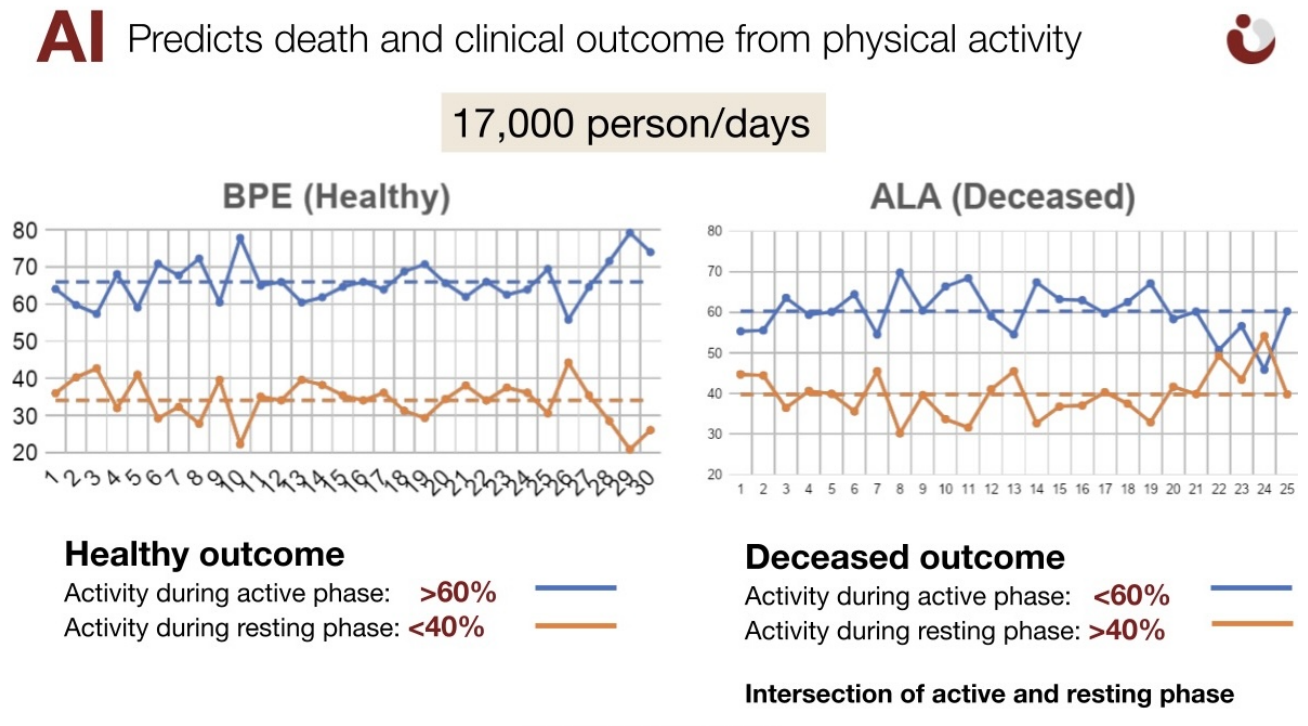
$$\text{Daily active percentage difference (\%)} = \text{Active percentageDaytime} - \text{Active percentageNighttime} \quad (10)$$

ML Algorithm for PA Analysis

Three key features—daily ELD, normalized ELD, and daily active percentage difference—were used to create the algorithm to predict the possible clinical worsening of discharged patients who demonstrate specific PA patterns. A risk alert was generated when the values of the features that were lower than a specific threshold were detected. The collected data correlated with the detailed measurement of the patients' PA. We assessed the patients' quality of recovery through accurate measurements of their activities while they were awake, while performing various activities of daily living (ADL) or participating in PA, and while

resting. Multiple layers of big data analytics, data mining algorithms, and ML methods were tested. Specifically, we applied the XGBoost ML algorithm. Our solution refined the predictive capability by using the individual PA differences during the active phase (walking, standing, or sitting) versus the resting phase (lying down). XGBoost enhanced the predictive accuracy of healthy recovery versus deterioration at home and determined the need to contact health care professionals. XGBoost distinguished itself from other gradient boost learning methods by using clever penalization of trees, proportional shrinking of leaf nodes, Newton boosting, extra randomization parameter, and the implementation of single distributed systems. These features enabled efficient ML classification of the real-time monitoring of PA to refine the patients' risk assessment. XGBoost distributed the feed-forward module of PA. The integration started with the PA module of 3 physical acceleration features (1 g, 2 g, and 3 g). The penalty-based system determined the initial risk profiling by PA weight and retrained in a stage-agnostic way to determine the features and penalties to enhance the weight from PA. The final stage repeats the same cycle as stage 2 and provides each patient's final weight and risk score. As the individual stage is algorithm agnostic, this method provides randomized nonbiased Newtonian analysis. The training data set was achieved by our solution by predicting the clinical outcome based on the individual PA differences during the active phase (walking, standing, or sitting) versus the resting phase (lying down). On the basis of >17,000 person-day data points of 36 participants (unpublished data), we predicted healthy recovery versus death in skilled nursing facility residents based on the PA data and analysis (Figure 2). Using this training data set, our XGBoost algorithm was designed to detect deterioration in the health condition of the discharged patients to generate a risk alert, which suggested the need for early medical intervention by contacting health care professionals, and prevent hospital readmission.

Figure 2. Data for clinical prediction. (A) Healthy outcome (physical activity [PA] ratio of active and resting phase: >60/40 ratio). (B) Deteriorating (deceased) outcome (PA ratio of active and resting phase: <60/40 ratio). ALA and BPE are the anonymized names of the patients.



Statistics and the AUC-ROC Curve

The ACT-I trial was an observational study, and we performed a correlational analysis between PA and clinical outcomes by using the AI model. The population consisted of 249 patients discharged from Stanford Hospital general cardiology and heart failure services. The measurement units relate to 36 (14.5%) of the 249 patients who completed a 28-day analysis. The response is the clinical outcome, and the factor is a comprehensive profile of PA from the patients. The choice of model is XGBoost. XGBoost-generated graph is a commonly used graph that summarizes the performance of a classifier over all possible thresholds. It is generated by plotting the true-positive rate (y-axis) against the false-positive rate (x-axis) as the threshold for assigning observations to a given class. AUC measures the entire 2D area under the ROC curve. The maximum value it can reach is 1; generally, the greater the value, the better the performance of the model.

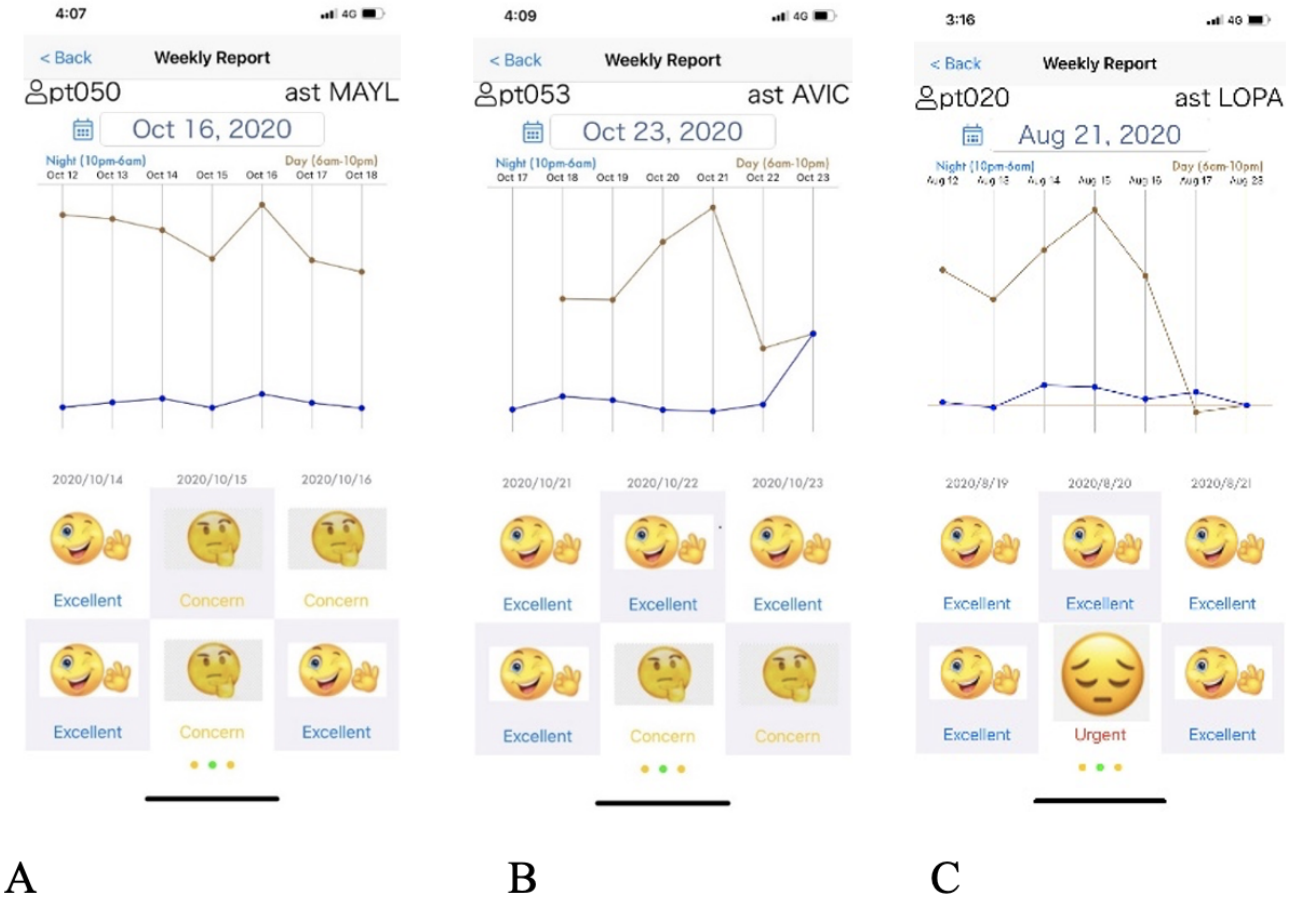
ROC analysis was used to evaluate a classifier's prediction performance in biological and medical applications. Each data point in the ROC curve comprises a pair: true-positive rate (sensitivity) and false-positive rate (1−specificity), generated by a discrete classifier with a specific threshold [14]. This study used several periods of MA from 3 days to produce all ROC points in the ROC space. Considering the effect of the imbalanced data set, meaning that the number of healthy discharged patients is greater than the number of discharged

patients classified as deteriorating, we also used the recall-precision curve to evaluate the performance of the proposed algorithms. In the recall-precision curve space, the x-axis and y-axis represent the recall values and precision values, respectively, calculated from the different thresholds.

User Interface

On the weekly report, we displayed the line chart, which kept track of the patient's activity level, and at the bottom, we generated a user-friendly emoji to report the health condition measured by our algorithm every 12 hours. The descriptor "Excellent" and a smiley face emoji indicate a healthy and normal pattern. The descriptor "Concern" and a pensive face emoji mean that there was an unusual pattern, indicating that the patients should be aware of the possible worsening of their clinical condition. Finally, the descriptor "Urgent" and a sad face emoji signify an unhealthy signal from the patient's PA pattern. This prediction suggests that the patients should contact their health care professionals (Figure 3).

Figure 3. Graphic user interface for alert notifications in 3 representative patients. MAYL, AVIC, and LOPA are also anonymized names of patients.The descriptor and smiley face emoji indicate (A) "Excellent" health followed by the descriptor "Concern" and pensive face emoji for mild risk, (B) "Excellent" health followed by the descriptor "Concern" and pensive face emoji for mild risk, and (C) "Excellent" health followed by the descriptor "Urgent" and sad face emoji for indication of risk. ast: assistance.



Results

Patient Enrollment

We screened 249 patients discharged from Stanford Hospital general cardiology and heart failure services. Of these 249 patients, 52 (20.9%) were enrolled, and 36 (14.5%) completed a 28-day analysis. The reasons for noncompletion were as follows: (1) withdrawal from study (10/16, 63%), (2) battery failure (4/16, 25%), and (3) early readmission (2/16, 13%). Of the 36 patients, 30 (83%) responded to the ADL questionnaire and 30 (83%) responded to the satisfaction questionnaire.

Prediction

Our data demonstrated a robust prediction system to forecast the worsening clinical outcomes of these patients based on their PA data, achieving a sensitivity of 87% and a specificity of 79%. On the basis of real-time assessment of PA, our technology offered clinically reliable predictions regarding the discharged patients who would need to contact their health care professionals or caretakers to report their worsening clinical condition. This capability allowed early intervention to prevent further deterioration of these patients (Figure 4) [10].

After the patient's discharge to home, the AiCare platform is deployed to the patient via the Rockband and iPhone. PA data and trend are displayed on the iPhone app. If there is any negative change in PA, the patient is contacted for a clinical evaluation. The ADL questionnaire is administered to the patient to see whether there is any correlation.

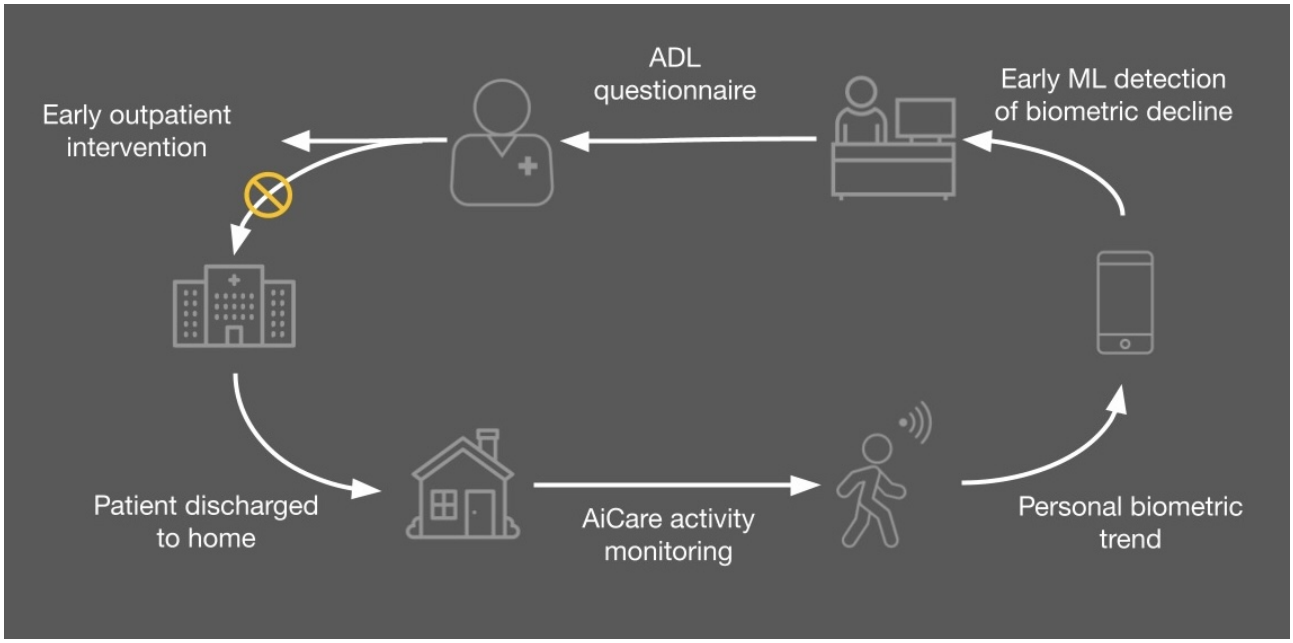
Our solution provided a robust low-cost technology to measure PA and predict clinical outcomes. Our platform also included the nudge technology for an AI-enabled questionnaire. This platform was designed to deliver a seamless, low-cost, and user-friendly environment for the remote monitoring of discharged patients at home to empower the patients and family members. This study analyzed the predictive capability of our platform as described in Textboxes 2-4.

The clinical diagnoses were categorized into true positive, true negative, false positive, and false negative as shown in Textbox 5.

The technical parameters were achieved and categorized into true positive (TP), true negative (TN), false positive (FP), and false negative (FN) as shown in Textbox 6.

Questionnaires were also administered at the time of risk alert as well as at the completion of the study (Textbox 7). The ADL questionnaire was administered to augment the PA data. The findings demonstrated modest correlation with the predictive capability. The patients whose condition deteriorated (true-positive group) showed the lowest function in terms of ADL, whereas those who remained stable showed higher response scores (true-negative group). However, there were low ADL scores in the false-positive group and high ADL scores in the false-negative group. The responses to the satisfaction questionnaire demonstrated that this platform was well received. The majority of the users stated that they would recommend the technology to others.

Figure 4. Patient flow. ADL: activities of daily living; ML: machine learning.



Textbox 2. Duration to risk prediction and intervention categorized into true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

- Mean number of days from discharge to risk prediction: TP=9 (SD 4); TN=none; FP=7 (SD 3); and FN=none
- Mean number of days from risk prediction to patient-initiated contact of health care professional: TP=5 (SD 2); TN=none; FP=2 (SD 2); and FN=none
- Total activity (1 g, 2 g, and 3 g per patient): TP=33,351 (SD 15,774); TN=38,998 (SD 19,062); FP=43,430 (SD 16,638); and FN=30,714 (SD 16,998)

Textbox 3. Formulas of area under the receiver operating characteristic curve metrics (TP=true positive, TN=true negative, FP=false positive, and FN=false negative).

Accuracy:

Precision and positive predictive values:

Sensitivity, recall, or true-positive rate (TPR):

Specificity or true-negative rate:

Negative predictive values (NPV):

False-positive rate (FPR)= 1 − specificity:

Textbox 4. Calculated values for the area under the receiver operating characteristic curve.

- Sensitivity: 90.01%
- Specificity: 81.55%
- Positive predictive value: 77.1%
- Negative predictive value: 92.3%
- Accuracy: 85%
- False-positive rate: 18.45%

Textbox 5. Clinical diagnosis and prediction data.

True positive (n=17)

- Heart failure (n=9)
- Arrhythmia (n=6)
- Atrial (n=5)
- Ventricular (n=1)
- Device (n=1)
- Ischemia (n=1)

True negative (n=9)

- Arrhythmia (n=6)
- Atrial (n=4)
- Ventricular (n=2)
- Ischemia (n=2)
- Heart failure (n=1)

False positive (n=11)

- Arrhythmia (n=5)
- Atrial (n=3)
- Ventricular (n=2)
- Heart failure (n=2)
- Ischemia (n=2)
- Pulmonary hypertension (n=2)

False negative (n=2)

- Arrhythmia, ventricular (n=1)
- Pericarditis (n=1)

Textbox 6. Technical findings.

- Signal loss hours per patient: TP=79; TN=74; FP=71; and FN=60
- Battery life per patient (d): TP=34; TN=29; FP=32; and FN=19
- Early replacement of the Rockband (number of patients): TP=1; TN=4; FP=0; and FN=1

Textbox 7. Response to the 7-question activities of daily living (ADL) questionnaire (7/7, 100%: highest function; n=30).

- True positive: 5.5 (positive ADL engagement)
- True negative: 6.25
- False positive: 4.7
- False negative: 7

Discussion

Principal Findings

We screened 249 patients discharged from Stanford Hospital general cardiology and heart failure services. Of these 249 patients, 52 (20.9%) were enrolled, and 36 (14.5%) completed a 28-day analysis looking into the correlation between PA and clinical outcome. Using our XGBoost model, we plotted the true-positive rates against false-positive rates, which helped us to generate the AUC-ROC curve and calculate the 2D AUC value to determine the performance of the model. Our data demonstrated a robust prediction system to forecast the worsening clinical outcomes of these patients based on their PA data, achieving a sensitivity of 87% and a specificity of 79%.

This innovative platform enables low-cost, robust, and precise PA tracking of patients discharged from hospital to predict stable versus unstable clinical recovery, using a WD and an iPhone. This technology is powered by our algorithms, individualized big data, personalized behavior- and function-specific web-based software, and intelligent ML analytics. This comprehensive platform offers an effective convergence of eHealth, AI, and telemedicine technology over internet-enabled mobile devices to leverage the economical, low-cost, and pervasive internet technology and, potentially, may address the socioeconomic divide seen today. Patient care at home by family members or by the individual patient is personalized by AI for maximum safety. This technology will fill an important gap in telemedicine through the use of user-friendly, real-time, and 24/7 remote monitoring for clinical outcome prediction. Patients in transition who are discharged from the hospital, emergency department, or urgent care clinic will benefit from this technology, which can monitor their progress and predict clinical deterioration to enable early intervention for successful recovery at home.

This ACT-I trial used monitoring technology to measure in-home activity and predict clinical outcomes. Although many innovative technologies claim accurate measurements of vital signs, there is no platform with proper validation of clinical outcome prediction data. A wide range of longitudinal studies to demonstrate the effectiveness of remote monitoring technology have been performed [10,15]. Most research was conducted within the area of passive infrared motion sensor technology, followed by research on body-worn sensors. Although the research into the use of monitoring technologies has been extensive, most studies only focused on demonstrating the functionality of the proposed monitoring technology by simulating activities in a laboratory setting or on an existing data set. As a result, the functionality of most systems has only been demonstrated in general terms or mechanical accuracy, sensitivity, and specificity. The long-term clinical effects of using monitoring technology are less well studied; for instance, in a meta-analysis on ambient sensors for older adult care, 25 of the 141 studies were pilot studies, with 11 focusing on the use of passive infrared motion sensor technology and 10 on the use of multicomponent monitoring technology. Study durations ranged from 3 weeks to 3 years [16]; only 4 studies were longitudinal, including 1 randomized controlled trial and 1 implementation study [17]; and all focused on the use of motion sensor technology. WDs have evolved from merely telling time to encompassing ubiquitous computing applications, miniaturized sensors, and wearable computer technology. Fitbit released its first wearable watch in 2009 and focused on activity tracking. During the ensuing years, smartwatches became common technology products manufactured by electronics companies. These developments led to a set of design guidelines for wearability and WDs that make tracking PA a much more attractive target for discharged patients [9].

PA is one of the major determinants for overall clinical outcomes in chronic diseases, including diabetes, hypertension, and heart disease, as well as mental health issues [17]. Reaching a sufficient level of PA could reduce the risk of cardiovascular disease (CVD), type 2 diabetes, obesity, depression, and anxiety [18-20]. PA even plays an important role in cancer prevention at specific sites, including breast and colon cancers [16]. In 2017, the Centers for Disease Control and Prevention advocated adding PA as 1 of the 4 vital signs [20]. Despite these efforts, automated clinical outcome prediction systems do not exist. There is a need for the accurate prediction of morbidity and mortality, particularly among older adults who are most frequently readmitted to the hospital. Patients with chronic diseases, cardiometabolic syndrome, and dementia are often underserved, growing in numbers, incurring higher costs to our society, and becoming increasingly vulnerable. Our large data set obtained from a skilled nursing facility and consisting of >17,000 person-day data points (408,000 person-hour data points of 36 patients captured over 2 years), demonstrated a high correlation of PA analysis among the residents who survived versus those who died. Using this as our training data set, we were able to identify with high accuracy patients who experienced stable recovery versus those who experienced unstable recovery during the most vulnerable 1-month posthospital discharge period.

The evidence-based management of CVDs requires substantial amounts of resources, including advanced therapeutics, complex diagnostics, and sophisticated clinical trials. However, the reliable prediction of clinical outcomes after hospital discharge has presented some challenges [11]. In response, various approaches using ML models such as artificial neural network, decision tree, support vector machine, and naïve Bayes have been attempted to predict clinical outcomes, taking into account steps, vital signs, medical conditions, and demographic information [11]. In one of the studies conducted on arrhythmic sudden cardiac death, a deep learning technology approach termed Survival Study of Cardiac Arrhythmia Risk was developed to predict risk for 156 patients with ischemic heart disease. In this model, cardiac magnetic resonance images and covariate data such as demographics, risk factors, electrocardiogram (ECG) measurements, medication use, and outcomes were used as inputs for the 2 branches, where 1 branch is used to visualize the heart's 3D ventricular geometry, and the other is used to extract arrhythmic sudden cardiac death risk–related imaging features from the cardiac magnetic resonance images. All these data were then used to create a survival curve individualized for each patient with accurate predictions for up to 10 years. However, the limitations of this study include an inability to account for competing risks for the same symptoms and covariates that were not exhaustive or fully inclusive [17]. In another instance, regression and convolutional neural network models were used to predict CVD risk for women. As CVDs are the primary cause of death in women, with evidence of sex bias in the diagnosis of CVDs , the exploration of screening factors for risk detection has never been more urgent. This study assessed the critical risk–screening opportunities offered to women and how the integration of AI can greatly benefit health care providers in interpreting data on women [17,21]. AI may propel the analysis of patient data into meaningful interpretations of patient health, providing health care providers with an additional layer of guidance for patient management plans. After patient discharge, ML is still viable in assisting with remote health monitoring through systems such as Wanda-CVD, which uses patients' blood pressure and BMI

measurements as well as their low-density lipoprotein and high-density lipoprotein cholesterol levels to coach them and improve their risk factors for CVD. However, using limited inputs such as blood pressure readings and cholesterol levels may not be entirely meaningful. In a previous study, only less than half of the predictions based solely on cholesterol levels and BMI measurements were correct [8,22].

In contrast, our XGBoost ML model enhanced the accuracy of predicting stable recovery versus clinical deterioration after discharge from the hospital. Specifically, XGBoost performs self-adaptive feature selection and prioritization among our data dimensions, mitigates the risk of overfitting by controlling the complexity of the trees with penalization on leaf nodes to cope with the high-frequency nature of our temporal data set, uses a Newton boosting algorithm to better learn the tree structures, and decorrelates the individual trees with a randomization parameter to reduce the bias and variance of the model. Compared with the existing approach, our AI-enabled solution is unique in the following ways. First, our real-time PA-based algorithm enables risk prediction during the critical 1-month posthospital discharge period, whereas most of the other research on risk prediction focused on a much longer time frame of 5 to 10 years. Our approach allows a shift to early intervention and the prevention of clinical deterioration during the postdischarge period. Second, our training data set for the ACT-I clinical trial consisted of a large number of data sets obtained during a long follow-up period. Our data covered a 2-year duration, which enabled a longitudinal follow-up for a personalized benchmark to conduct individualized analysis and risk prediction. Third and last, our low-cost at-home patient onboarding process did not rely on complex hardware such as imaging or remote ECG equipment. The Rockband WD was low cost, maintenance free, and disposable. Our hardware-agnostic AI framework demonstrated highly and easily adaptable features using our simple WD.

Limitations

Although the majority of the patients were satisfied with our platform, there were some compliance issues related to the use of the WD. Furthermore, the measurement of PA only may not provide a comprehensive assessment and prediction of an individual patient's clinical condition. Our future trial, ACT-II, will expand on the ACT-I trial's limitations by improving the specificity (false positive) rate of the ACT-I trial by evaluating the efficacy of an augmented XGBoost algorithm. We will use an Apple Watch to complement PA measurements by also monitoring HR, HR variability, ECG, oxygen saturation, blood pressure (separate blood pressure measurement device), clinical data, and genomics to better identify stable versus unstable recovery. Our novel platform in an iOS environment will enable the capture of multidimensional real-time data to enhance patients' awareness of their clinical condition and health care professionals' guidance of patient management. We will investigate the feasibility of this platform, consisting of an Apple Watch, an iPhone, an XGBoost interface, and a HIPAA-compliant AWS environment, to monitor the dynamic biometric data, predict patients' clinical outcome, and improve patient compliance.

Conclusions

The ACT-I trial demonstrated a critical proof of concept of the Rockband WD to enable real-time analysis of patients' PA data remotely. We developed a cloud-enabled XGBoost algorithm and intelligent sensor technology to enable precision home health care. The XGBoost algorithm quantified, integrated, and predicted the pattern of each patient's outcome seamlessly with high accuracy, precision, and recall. The Rockband cloud backend personalized the big data for behavior- and function-specific interactive software, and ML analytics allowed a comprehensive platform to converge eHealth, AI, and telemedicine technology. Our internet-enabled mobile devices leveraged the economical, low-cost, and pervasive technology to personalize health care by enabling prevention and early intervention through the real-life clinical implementation of mobile device technology and AI. Our approach developed, tested, and disseminated the next generation of health care strategy by focusing on precision health, using diagnostic information collected in real time from patients' PA data while they were recovering at home. Our XGBoost algorithm enabled this scalable, portable, and distributed processing framework. This novel technology will introduce a nascent approach to patient care to redefine clinical practice by predicting patient outcome based on a comprehensive analysis of behavioral phenotype. This real-time risk monitoring and clinical outcome prediction platform will advance the future of remote patient care.

Acknowledgments

The authors appreciate and acknowledge the assistance of Fouzia Khan and Banu Priya Rathinam Radha Rajasekaran in patient enrollment and trial organization.

Data Availability

We hope to submit our data to *JMIR Data*. However, we obtained the informed consent to release their data in our repository; therefore, this may not be possible.

Conflicts of Interest

The manuscript presents a potential conflict of interest due to the financial involvement of its authors, WX's family member, PCY, ZS, AJ, and PJ in AiCare, the company responsible for sponsoring the clinical trials discussed in the manuscript. The conflict arises from the authors holding shares in AiCare, indicating a direct financial interest in the success and promotion of

the company's products. Full disclosure and transparency regarding these financial relationships are essential for maintaining the integrity of the scientific work and ensuring the reader's ability to assess potential biases.

References

1. Golightly YM, Allen KD, Ambrose KR, Stiller JL, Evenson KR, Voisin C, et al. Physical activity as a vital sign: a systematic review. Prev Chronic Dis 2017 Nov 30;14:E123 [FREE Full text] [doi: 10.5888/pcd14.170030] [Medline: 29191260]
2. Walker B. The top 5 causes of hospital readmissions – and how to prevent them. Upfront Healthcare. 2022 Feb 16. URL: https://www.patientbond.com/blog/the-top-5-causes-of-hospital-readmissions-and-how-to-prevent-them [accessed 2022-08-24]
3. McIlvennan CK, Eapen ZJ, Allen LA. Hospital readmissions reduction program. Circulation 2015 May 19;131(20):1796-1803 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.114.010270] [Medline: 25986448]
4. Franklin BA, Wedig IJ, Sallis RE, Lavie CJ, Elmer SJ. Physical activity and cardiorespiratory fitness as modulators of health outcomes: a compelling research-based case presented to the medical community. Mayo Clin Proc 2023 Feb;98(2):316-331 [FREE Full text] [doi: 10.1016/j.mayocp.2022.09.011] [Medline: 36737120]
5. Yang CC, Hsu YL. Development of a portable system for physical activity assesement in a home enviorment. Feng Chia University Global Information System. URL: https://dspace.fcu.edu.tw/handle/2377/3692 [accessed 1999-11-30]
6. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: an accurate and reliable device for wireless physical activity tracking. Int J Cardiol 2015 Apr 15;185:138-140 [FREE Full text] [doi: 10.1016/j.ijcard.2015.03.038] [Medline: 25795203]
7. Parab S, Bhalerao S. Study designs. Int J Ayurveda Res 2010 Apr;1(2):128-131 [FREE Full text] [doi: 10.4103/0974-7788.64406] [Medline: 20814529]
8. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
9. Gemperle F, Kasabach C, Stivoric J, Bauer M, Martin R. Design for wearability. In: Proceedings of the 1998 Digest of Papers, 2nd International Symposium on Wearable Computers (Cat. No.98EX215). 1998 Presented at: SWC '98; October 19-20, 1998; Pittsburgh, PA p. 116-122 URL: https://ieeexplore.ieee.org/document/729537 [doi: 10.1109/iswc.1998.729537]
10. Mkrtchyan FA. On the effectiveness of remote monitoring systems. In: Proceedings of the 2018 Earth Observing Systems XXIII on Spie Optical Engineering and Applications. 2018 Presented at: EOS '18; August 19-23; San Diego, CA URL: https://doi.org/10.1117/12.2322522 [doi: 10.1117/12.2322522]
11. Katarya R, Srinivas P. Predicting heart disease at early stages using machine learning: a survey. In: Proceeedings of the 2020 International Conference on Electronics and Sustainable Communication Systems. 2020 Presented at: ICESC '20; April 28-30; Coimbatore, India p. 302-305 URL: https://ieeexplore.ieee.org/document/9155586 [doi: 10.1109/icesc48915.2020.9155586]
12. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Indian Pediatr 2011 Apr;48(4):277-287. [doi: 10.1007/s13312-011-0055-4] [Medline: 21532099]
13. Buchman AS, Boyle PA, Yu L, Shah RC, Wilson RS, Bennett DA. Total daily physical activity and the risk of AD and cognitive decline in older adults. Neurology 2012 Apr 24;78(17):1323-1329. [doi: 10.1212/wnl.0b013e3182535d35]
14. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006 Jun;27(8):861-874 [FREE Full text] [doi: 10.1016/j.patrec.2005.10.010]
15. Infeld M, Wahlberg K, Cicero J, Plante TB, Meagher S, Novelli A, et al. Effect of personalized accelerated pacing on quality of life, physical activity, and atrial fibrillation in patients with preclinical and overt heart failure with preserved ejection fraction: the myPACE randomized clinical trial. JAMA Cardiol 2023 Mar 01;8(3):213-221. [doi: 10.1001/jamacardio.2022.5320] [Medline: 36723919]
16. Uddin MZ, Khaksar W, Torresen J. Ambient sensors for elderly care and independent living: a survey. Sensors (Basel) 2018 Jun 25;18(7):2027 [FREE Full text] [doi: 10.3390/s18072027] [Medline: 29941804]
17. Steele BG, Holt L, Belza B, Ferris S, Lakshminaryan S, Buchner DM. Quantitating physical activity in COPD using a triaxial accelerometer. Chest 2000 May;117(5):1359-1367. [doi: 10.1378/chest.117.5.1359] [Medline: 10807823]
18. Brodie MA, Coppens MJ, Lord SR, Lovell NH, Gschwind YJ, Redmond SJ, et al. Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different. Med Biol Eng Comput 2016 Apr 6;54(4):663-674. [doi: 10.1007/s11517-015-1357-9] [Medline: 26245255]
19. McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, et al. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart counts cardiovascular health study. JAMA Cardiol 2017 Jan 01;2(1):67-76. [doi: 10.1001/jamacardio.2016.4395] [Medline: 27973671]
20. Popescu DM, Shade JK, Lai C, Aronis KN, Ouyang D, Moorthy MV, et al. Arrhythmic sudden death survival prediction using deep learning analysis of scarring in the heart. Nat Cardiovasc Res 2022 Apr;1(4):334-343 [FREE Full text] [doi: 10.1038/s44161-022-00041-9] [Medline: 35464150]
21. Adedinsewo DA, Pollak AW, Phillips SD, Smith TL, Svatikova A, Hayes SN, et al. Cardiovascular disease screening in women: leveraging artificial intelligence and digital tools. Circ Res 2022 Feb 18;130(4):673-690 [FREE Full text] [doi: 10.1161/CIRCRESAHA.121.319876] [Medline: 35175849]

22. Alshurafa N, Eastwood JA, Pourhomayoun M, Liu JJ, Sarrafzadeh M. Remote health monitoring: predicting outcome success based on contextual features for cardiovascular disease. Annu Int Conf IEEE Eng Med Biol Soc 2014;2014:1777-1781. [doi: 10.1109/EMBC.2014.6943953] [Medline: 25570321]

Abbreviations

ACT-I: Activity Assessment of Patients Discharged from Hospital-I
ADL: activities of daily living
AI: artificial intelligence
API: application programming interface
AUC: area under the curve
AUC-ROC: area under the receiver operating characteristic curve
AWS: Amazon Web Services
BLE: Bluetooth low energy
CVD: cardiovascular disease
ECG: electrocardiogram
EL: energy level
ELD: energy level difference
HIPAA: Health Insurance Portability and Accountability Act
HR: heart rate
HRRP: Hospital Readmissions Reduction Program
MA: moving average
ML: machine learning
PA: physical activity
ROC: receiver operating characteristic
SMA: simple moving average
SVM: signal vector magnitude
WD: wearable device
XGBoost: extreme gradient boosting

Edited by A Mavragani; submitted 16.12.22; peer-reviewed by S Sarejloo, K Gupta, T Bonten, X Cheng; comments to author 09.05.23; revised version received 31.08.23; accepted 19.09.23; published 01.03.24.

Please cite as:
Yang PC, Jha A, Xu W, Song Z, Jamp P, Teuteberg JJ
Cloud-Based Machine Learning Platform to Predict Clinical Outcomes at Home for Patients With Cardiovascular Conditions Discharged From Hospital: Clinical Trial
JMIR Cardio 2024;8:e45130
URL: https://cardio.jmir.org/2024/1/e45130
doi:10.2196/45130
PMID:38427393

Original Paper

Identifying Predictors of Heart Failure Readmission in Patients From a Statutory Health Insurance Database: Retrospective Machine Learning Study

Rebecca T Levinson[1*], PhD; Cinara Paul[1*], Dr med; Andreas D Meid[2], Dr sc hum; Jobst-Hendrik Schultz[1], Dr med; Beate Wild[1], Dr sc hum

[1]Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg University, Heidelberg, Germany

[2]Medical Faculty of Heidelberg, Internal Medicine IX - Department of Clinical Pharmacology and Pharmacoepidemiology, Heidelberg University Hospital, Heidelberg University, Heidelberg, Germany

[*]these authors contributed equally

Corresponding Author:
Rebecca T Levinson, PhD
Department of General Internal Medicine and Psychosomatics
Heidelberg University Hospital
Heidelberg University
Im Neuenheimer Feld 410
Heidelberg, 69120
Germany
Phone: 49 6221565888
Email: rebeccaterrall.levinson@med.uni-heidelberg.de

Abstract

Background: Patients with heart failure (HF) are the most commonly readmitted group of adult patients in Germany. Most patients with HF are readmitted for noncardiovascular reasons. Understanding the relevance of HF management outside the hospital setting is critical to understanding HF and factors that lead to readmission. Application of machine learning (ML) on data from statutory health insurance (SHI) allows the evaluation of large longitudinal data sets representative of the general population to support clinical decision-making.

Objective: This study aims to evaluate the ability of ML methods to predict 1-year all-cause and HF-specific readmission after initial HF-related admission of patients with HF in outpatient SHI data and identify important predictors.

Methods: We identified individuals with HF using outpatient data from 2012 to 2018 from the AOK Baden-Württemberg SHI in Germany. We then trained and applied regression and ML algorithms to predict the first all-cause and HF-specific readmission in the year after the first admission for HF. We fitted a random forest, an elastic net, a stepwise regression, and a logistic regression to predict readmission by using diagnosis codes, drug exposures, demographics (age, sex, nationality, and type of coverage within SHI), degree of rurality for residence, and participation in disease management programs for common chronic conditions (diabetes mellitus type 1 and 2, breast cancer, chronic obstructive pulmonary disease, and coronary heart disease). We then evaluated the predictors of HF readmission according to their importance and direction to predict readmission.

Results: Our final data set consisted of 97,529 individuals with HF, and 78,044 (80%) were readmitted within the observation period. Of the tested modeling approaches, the random forest approach best predicted 1-year all-cause and HF-specific readmission with a C-statistic of 0.68 and 0.69, respectively. Important predictors for 1-year all-cause readmission included prescription of pantoprazole, chronic obstructive pulmonary disease, atherosclerosis, sex, rurality, and participation in disease management programs for type 2 diabetes mellitus and coronary heart disease. Relevant features for HF-specific readmission included a large number of canonical HF comorbidities.

Conclusions: While many of the predictors we identified were known to be relevant comorbidities for HF, we also uncovered several novel associations. Disease management programs have widely been shown to be effective at managing chronic disease; however, our results indicate that in the short term they may be useful for targeting patients with HF with comorbidity at increased risk of readmission. Our results also show that living in a more rural location increases the risk of readmission. Overall, factors beyond comorbid disease were relevant for risk of HF readmission. This finding may impact how outpatient physicians identify and monitor patients at risk of HF readmission.

(JMIR Cardio 2024;8:e54994) doi:10.2196/54994

KEYWORDS

statutory health insurance; readmission; machine learning; heart failure; heart; cardiology; cardiac; hospitalization; insurance; predict; predictive; prediction; predictions; predictor; predictors; all cause

Introduction

Patients with heart failure (HF) are the most commonly readmitted group of adult patients in Germany and other Western industrialized countries [1,2]. Nearly two-thirds of patients with HF are readmitted within 1 year [3]. Accounting for ~1%-2% of the annual health care expenditure, with roughly 60% of the spending attributed to inpatient stays, HF poses a major economic burden for health systems, particularly for those who offer universal health coverage [4]. Besides, readmissions increase the risk of complications and mortality in patients with HF [5]. Therefore, understanding the contributors to readmission for identifying patients at risk would be a major step toward both the improvement of patient care and the reduction of costs associated with HF.

Most studies for prediction of HF readmission are based on data from trials and electronic health records introducing a risk for selection bias [6]. Routinely collected data from statutory health insurance (SHI) companies provide large longitudinal data sets representative of the general population. The advantages include reflecting comprehensive and real-life health care provisions for all insured people [7]. Health insurance is mandatory in Germany, with about 90% of the population having SHI [8]. Membership is open to everyone, independent of income, age, or state of health [9].

Outpatient data can provide a different window into the disease state, for example, outpatient data are known to capture a broader spectrum of comorbidity than may be present in inpatient data alone [10]. This may be crucial to the early identification of individuals at risk of readmission for noncardiovascular reasons in this patient group [11]. Furthermore, understanding the relevance of HF management outside the hospital setting is critical to understanding HF and the factors that lead to readmission [12]. To lower costs and ameliorate the patient's experience, understanding what noninvasive pathways within regular care should be targeted is vital.

To analyze large databases—such as SHI data—machine learning (ML) algorithms are promising methods. ML algorithms can process big data and identify complex patterns while being able to build both linear and nonlinear models for the association between predictor variables and outcomes [13]. ML techniques in cardiovascular research are an emerging field that may offer support in clinical decision-making [14]. ML approaches have successfully been implemented to predict coronary artery disease and atrial fibrillation [15,16]. A recent review concluded that ML algorithms had better discrimination than conventional statistical methods in predicting readmission risk in HF [17]. A recently published study from the Netherlands [18] investigated the predictors of HF-specific readmission using ML on SHI data. However, most readmissions in patients with HF are for noncardiovascular reasons, such as renal failure or pneumonia [19]. To the best of our knowledge, to date, no study exists that applied ML to only outpatient SHI data to predict all-cause readmission in HF.

The aims of this study were (1) to evaluate the use of outpatient SHI data to predict 1-year all-cause (primary end point) and HF-specific (secondary end point) readmission after an initial admission for HF and (2) to identify and rank relevant predictors for readmission. In order to target patients who are at-risk at the earliest possible stage, we included patients with HF who were hospitalized for the first time for HF and thus were just at the presumed start of the "HF readmission circle."

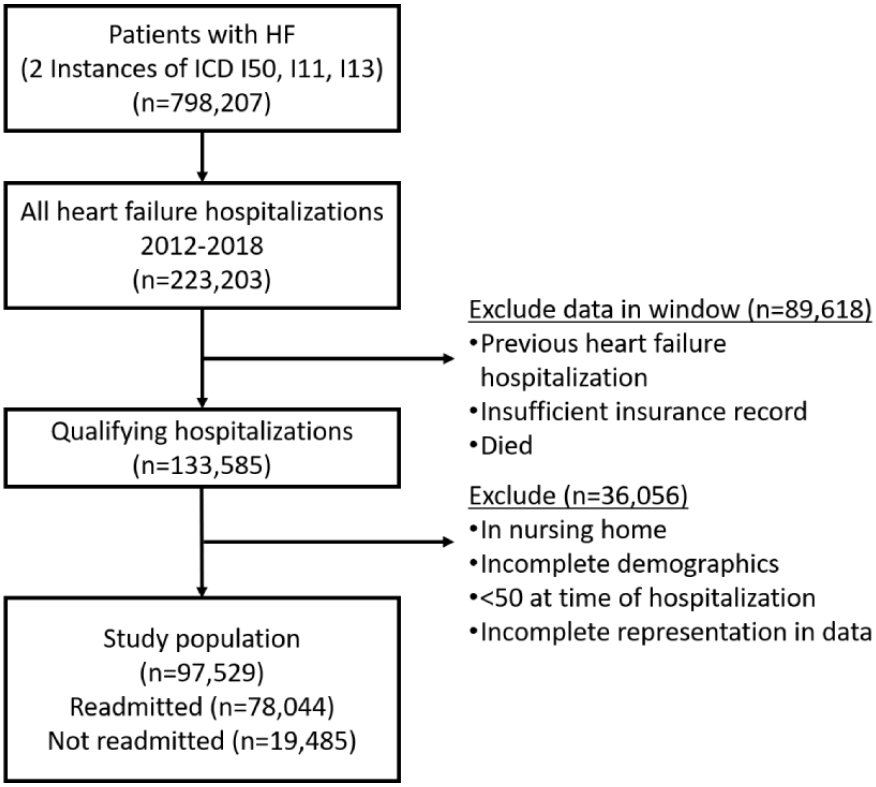
Methods

Study Population

We obtained anonymized data from health insurance claims (from 2012 to 2018) provided by the AOK Baden-Württemberg, a large German SHI with about 4.5 million insured people. In Germany, about 90% of the population receives coverage by SHI, of which the AOK overall company comprises >30% [8]. Within Baden-Württemberg, where the data used in this study originated, AOK comprises 45.5% of the population covered by SHI.

We included patients who had HF as documented by 2 or more instances of the *International Classification of Disease, 10th Revision* (*ICD-10*) code I50*, I13*, or I11* on either inpatient or outpatient records and on at least 2 different days. Figure 1 shows the sample selection process. To ensure that patients with 1 readmission were not being compared to those with many, we identified individuals who had their first HF-related admission from 2013 to 2017. All hospital stays were determined from hospital stay data. Hospital stays with shared dates were merged into 1, and at least 3 days were required between the end of the primary HF hospital stay and a potential readmission. To obtain admissions due to HF, *ICD-10* codes documenting reason for inpatient care were mapped to patient stay data. Individuals were required to have a year of record prior to their first HF admission to increase the likelihood of finding the first HF admission for a patient. Individuals were also required to have a year of record after their HF admission, unless they were readmitted. Individuals missing demographics, including date of birth and sex, and those who had insufficient insurance record during the observation period were also excluded. For the remaining population, age at HF diagnosis was calculated and those younger than 50 years at HF diagnosis or who lived in a nursing home were excluded from modeling.

Figure 1. Flowchart for identification of the study population. Patients were identified within statutory health insurance data (2012-2018) from AOK Baden-Württemberg, Germany. *ICD*: *International Classification of Disease*.



Study Outcomes

The primary end point of our study was first all-cause readmission within a year after an HF admission. To identify this, all admissions following the first HF admission, deemed the "index admission" in record, were identified. Patients with an all-cause admission 3-365 days after their index admission were considered to have been readmitted. Patients who did not have a readmission within 365 days but were alive and present in the data set on or after the 365-day mark were considered to not have been readmitted within the 1-year window. Patients who died or otherwise withdrew from the insurance scheme prior to the end of the 365 day within and who did not have a readmission were excluded from analysis.

As a secondary end point, we also evaluated the first readmission for HF after the index HF admission. The same methodology as for the primary end point was applied including the time frame of 1 year for readmission. However, readmissions were required to have an *ICD-10* code of I50*, I13*, or I11* attached to them to be considered HF specific.

Feature Curation and Selection for Prediction Models

To evaluate the role of comorbidities in the prediction of HF rehospitalization, *ICD-10* codes were obtained for all individuals in the study population. Codes were curated to remove entries that did not correspond to an *ICD-10* code and those with dates misentered to be outside the documented time period. Codes after the date of the first hospitalization were also excluded from analysis. *ICD-10* codes were then rolled up into their root code (eg, I25.1 and I25.2 both became I25). Codes on the same day were compiled into 1, and for each individual, the number of unique days each aggregated code appeared was counted. Codes that were part of the Z class of codes, indicating factors relevant to health care use, were also excluded from the analysis. The remaining codes were included as potential features for models. Medications were extracted from prescription medication documentation based on the Anatomical Therapeutic Chemical Classification index (ATC) assigned to each drug. For each ATC number, the number of total packages of a drug was multiplied by defined daily dose to estimate the cumulative in record exposure of an individual to a given drug. Drugs were then filtered to those ATC numbers representing the "C" class of drugs, those affecting the cardiovascular and circulatory system. For each drug, the estimated within data set exposure was included as a potential feature.

Demographic data were also obtained for individuals in the study population. Age, was calculated from date of birth and date of first HF hospitalization, and sex was included as likely relevant to clinical outcomes. Other demographic and demographic-derived variables (described in the following sentences) were included to account for socioeconomic status, professional status, and level of ease of access within the SHI. As we hypothesized that foreigners might have a different relationship with the German insurance system than a German national, a dichotomous variable indicating German nationality was included as a potential predictor. The type of coverage within the SHI was included as a variable with 3 levels, indicating primary holder, family insurance (as the spouse or other dependent family member of the primary insurance holder), or pensioner insurance. To account for a potential disparity in outcomes based on geography (a proxy for both wealth and access to hospitals), data indicating the degree of rurality for each administrative area in Germany were downloaded from the Thuenen Landatlas sponsored by the

German Ministry of Food and Agriculture [8,20]. This degree of rurality data was then mapped to the postal codes available in the AOK set, allowing evaluation of degree of rurality in our models. Participation in a disease management program (DMP) focused on diabetes mellitus type 1, diabetes mellitus type 2, breast cancer, chronic obstructive pulmonary disease (COPD), or coronary heart disease prior to the index HF admission was also obtained and included as a binary variable. Measures of cardiac structure or function, such as the output of electrocardiograms, echocardiography, or cardiac imaging, were not available in the data set, and therefore were not included in the prediction models.

Statistical Analyses

The study population was randomly split with a 70:30 ratio into a training and a testing set for modeling. Using individuals from the study population, 4 models were built for each end point: a logistic regression model, a stepwise regression, an elastic net, and a random forest (RF) model. For each model, potential features with nonmissing data in at least 99% of the training population were included, resulting in 265 features for potential inclusion. As the end point, readmission was unbalanced in the data set, subsampling with 10-fold cross-validation was used to reduce the bias toward predicting only rehospitalization. Elastic net was performed with 5-fold cross-validation, and admission outcomes were weighted based on their prevalence in the data set. Elastic net hyperparameters were tuned using a grid search with an α from 0 to 1 and a λ from 0.0001 to 2. For the RF model, the training set was used for model training and hyperparameter optimization using 3-fold cross-validation. Hyperparameter optimization was performed allowing between 50 and 500 trees, 3 and 20 nodes per tree, and 10 and 50 splits per node. Training was then performed to generate probabilities of readmission for each individual. Within the training set, a cut-point for prediction of readmission was then identified. The training model and cut-point were then evaluated in the testing data. Feature importance indicating the change in model performance due to the exclusion of variables was then generated from the final model. All predictors provided to the elastic net or RF were also included in a logistic regression model and provided in a backwards stepwise regression model. The model then used Akaike information criteria to reduce these features to the minimal set that best predicted HF.

For each modeling approach a C-statistic for model fit was calculated. For models that selected features, important features as determined by mean misclassification error rate through permutation were evaluated. All data management, modeling, and statistical analysis were performed with R (version 3.6.0, 2019-04-06; R Foundation for Statistical Computing) [21]. The packages *tidyverse* [22], *data.table* [23], *ggplot2* [24], *mlr3* [25], *caret* [25,26], and *pROC* [27] were used. For generation of tables summarizing demographics, chi-square tests or Wilcoxon rank sum tests were used as appropriate.

Ethical Considerations

This work was exempt from specific ethics approval as a secondary analysis of anonymized data (section 303e) [28]. In Germany, analyses of anonymized health insurance data do not require ethics committee approval by law.

Results

Population Characteristics

The final sample consisted of 97,529 patients with HF, with a median (IQR) age of 79 (70-85) years and an equal proportion of men (n=49,058, 50.3%) and women (n=48,471, 49.7%). Among them, 78,044 (80%) of the final sample were readmitted to the hospital within the observation period, but only 42,694 (43.2%) were readmitted with HF as one of the primary or secondary diagnoses. Table 1 summarizes baseline characteristics for the final sample and comparisons between those readmitted and not. Overall, readmitted patients were more likely to have pensioners insurance, lived in a more rural location, and had higher rates of outpatient codes for myocardial infarction and COPD. Comparisons between training and testing set are found in Table S1 in Multimedia Appendix 1 and readmission for HF-specific reasons can be found in Table S2 in Multimedia Appendix 1.

Individuals who were readmitted within a year after their initial HF hospitalization were often readmitted quickly, with 38% (n=29,747) of readmitted patients returning to the hospital within 30 days, 62% (n=48,628) within 90 days, and 78% (n=60,667) within 180 days (Figure 2A). For the HF-specific readmission end point, although a substantially smaller proportion of the population was readmitted, the trend for time to readmission was similar, with 70% (n=29,896) of readmitted patients readmitted within 180 days (Figure 2B).

Table 1. Demographics of the heart failure study population, stratified by all-cause readmission status within the observation period (2012-2018).

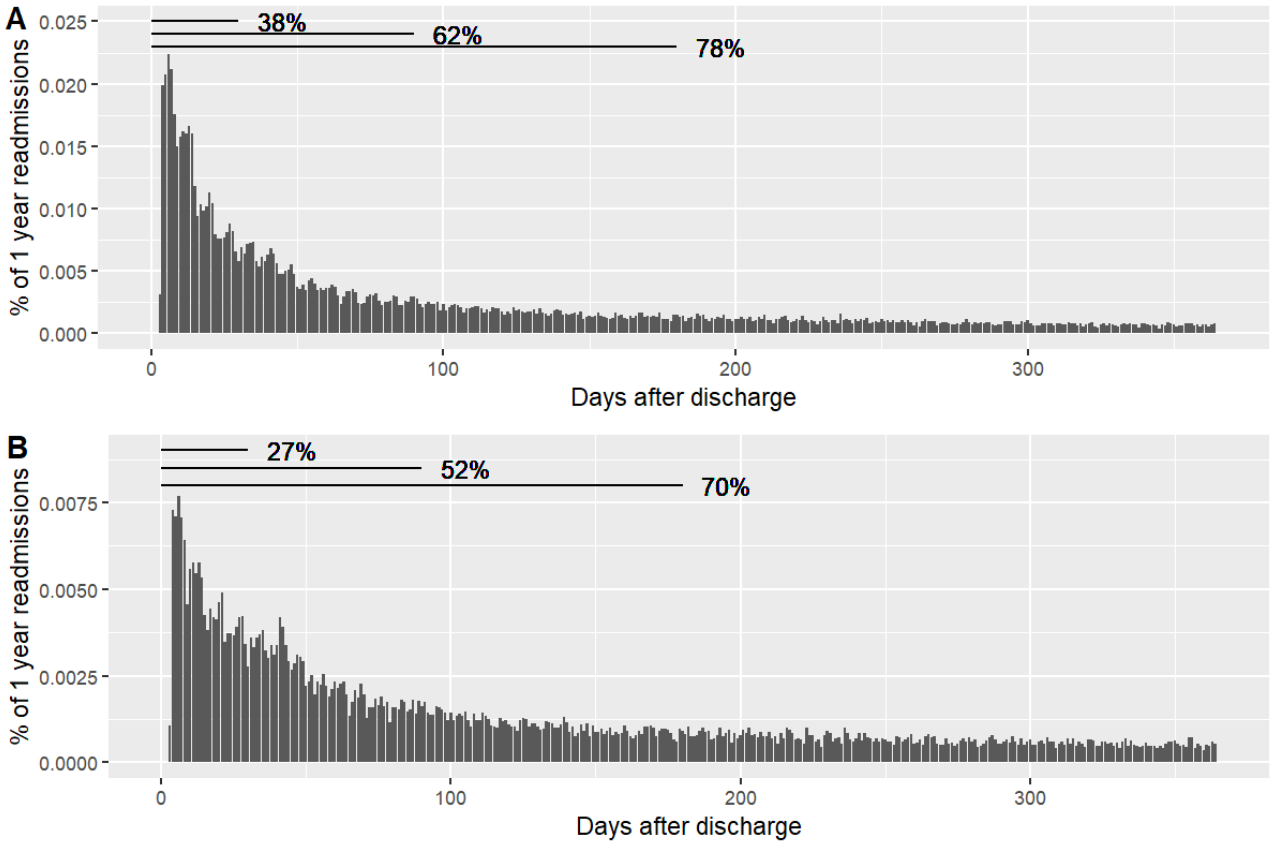
	All (N=97,529)	Readmitted (n=78,044)	Not readmitted (n=19,485)	*P* value[a]
Age (years), median (IQR)	79 (70 to 85)	79 (71 to 85)	79 (70 to 85)	.91
Sex (male), n (%)	49,058 (50)	40,237 (52)	8821 (45)	<.001
German national, n (%)	88,249 (90)	70,642 (91)	17,607 (90)	.45
Insurance type, n (%)				<.001
Primary holder	18,822 (19)	14,725 (19)	4097 (21)	
Family insurance	2231 (2)	1731 (2)	500 (3)	
Pensioner's insurance	76,476 (78)	61,588 (79)	14,888 (76)	
Degree of rurality, median (IQR)	0.06 (–0.52 to 0.53)	0.06 (–0.52 to 0.53)	0.08 (–0.52 to 0.54)	.01
Hypertension, n (%)	82,198 (84)	65,990 (85)	16,208 (83)	.13
Atrial fibrillation, n (%)	24,707 (25)	20,340 (26)	4367 (22)	.92
CAD[b], n (%)	41,384 (42)	33,942 (43)	7442 (38)	.31
Myocardial infarction, n (%)	7879 (8)	6612 (8)	1267 (7)	<.001
Hyperlipidemia, n (%)	51,415 (53)	41,442 (53)	9973 (51)	.84
Diabetes mellitus type 2, n (%)	41,342 (42)	33,919 (43)	7423 (38)	.71
COPD[c], n (%)	20,158 (21)	17,109 (22)	3049 (16)	<.001

[a]*P* values calculated based on chi-square or Wilcoxon rank sum tests as appropriate.

[b]CAD: coronary artery disease.

[c]COPD: chronic obstructive pulmonary disease.

Figure 2. Histogram of time to readmission for readmitted heart failure (HF) patients within the (A) all-cause and (B) HF-specific readmission cohorts. Percentages indicate percentage of the readmitted population for either all-cause or HF-specific readmission.



Model Performance

The performance of different models for the prediction of first all-cause readmission and the first HF readmission are provided in Table 2. For both, the all-cause and HF-specific readmission end points, the RF model provided the best model fit, with a C-statistic of 0.68 and 0.69, respectively. However, for the HF-specific end point, the elastic net and RF performed very similarly.

Table 2. C-statistics for model fit for the 4 applied modeling approaches. Statistics are provided for prediction of 1-year all-cause (primary end point) and heart failure–specific (secondary end point) readmission after an initial admission for heart failure.

	Logistic regression	Stepwise regression	Elastic net	Random forest
All-cause readmission	0.55	0.63	0.65	0.68
Heart failure–specific readmission	0.56	0.65	0.67	0.69

Predictors of Readmission

As the RF was the best performing model, we evaluated the features with the largest feature importance. The RF feature importance provides a level of importance to the model, but not the direction of the association; therefore, univariate analyses and effect sizes from the elastic net were used to provide additional context.

The most important predictor for the all-cause readmission end point according to the RF model was prescription of pantoprazole (Table 3 and Figure S1 in Multimedia Appendix 1). Other highly relevant features included COPD, sex, diabetes mellitus, atherosclerosis, peripheral vascular disease, age, participation in the coronary heart disease or diabetes DMPs. These predictors included known risk factors for both HF and for general cardiovascular health. In contrast, drugs included in this list tended to be more relevant to general conditions or pain. Degree of rurality was also among the predictors that had an impact on the final model.

For the HF-specific readmission end point, the most important features were the number of times an HF *ICD-10* code had been documented in the medical record prior to the index hospitalization and year of birth (Table 3 and Figure S2 in Multimedia Appendix 1). Other important features included atrial fibrillation, insurance type, chronic back pain, hypertension, degree of rurality, and hyperlipidemia. Overall, the most important features for HF-specific readmission included the majority of the most known and studied HF risk factors and comorbidities. The most important medication for HF-specific readmission was furosemide, a loop diuretic used to treat edema in patients with HF. Sex was significantly less important for the HF-specific model than it is for the all-cause RF. Enrolled in a DMP for diabetes mellitus type 2 or coronary heart disease was also an important predictor in this model.

Table 3. Top predictors for 1-year all-cause readmission in patients with heart failure by feature importance from the random forest model.

Feature[a]	Mean misclassification error[b]
A02BC02—pantoprazole	0.05445356
J44—other chronic obstructive pulmonary disease	0.024890419
Sex	0.02361295
E14—diabetes mellitus unspecified	0.015379432
I70—atherosclerosis	0.012226194
I73—other peripheral vascular disease	0.011574568
Age	0.011145327
I25—chronic ischemic heart disease	0.007728684
DM_KHK—DMP[c] coronary heart disease	0.007057427
DM_DM2—DMP diabetes mellitus type 2	0.005368298
E11—diabetes mellitus type 2	0.005053326
B01AB05—enoxaparin	0.005003516
I10—essential hypertension	0.004787284
R03BB04—tiotropium bromide	0.004478758
B01AA04—phenprocoumon	0.004253736
N18—chronic kidney disease	0.00281131
N19—renal insufficiency not otherwise specified	0.002636976
H02AB06—prednisolone	0.002499561
I35—nonrheumatic aortic valve disorders	0.00248579
M48—other spondylopathies	0.002316437
Degree of rurality	0.002048345
DM_COPD[d]—DMP COPD	0.001441254
A02BC01—omeprazole	0.001138002

[a]Feature name as provided in the data set is listed in the first column, followed by added annotation information, 7-digit codes indicate ATC classifications, and 3-character labels are *ICD-10* codes.

[b]Mean misclassification error represents the change in model score when each variable is randomly permuted.

[c]DMP: disease management program.

[d]COPD: chronic obstructive pulmonary disease.

Discussion

Principal Findings

Based on routinely collected health insurance data from >90,000 patients with HF, we have shown that exclusively using outpatient data has clear value for predicting 1-year HF-specific and all-cause readmission.

Interestingly, the 30-day rate of readmission in our analysis was higher than those in the previous studies. We found that 38% (29,747/78,044) of patients were readmitted for any cause, and 27% (11,377/42,694) were readmitted for HF within 30 days. In the same data set, although using a different classification of HF, Ruff et al [2] found that 21% of patients with HF were readmitted for HF within 30 days. It could be that this discrepancy is due to the inclusion of additional years of data with a higher rate of readmission or a difference in study design. However, though high, the rate of readmission seen at 1 year within this population is not implausible, given that others have reported 1-year readmission rates of approximately 67% [3].

The predictive ability of our models is similar to estimates from other retrospective analyses in the real-world data. Van der Galiën et al [18] was able to predict 1-year HF readmission with a C-statistic of 0.71-0.73 including both inpatient and outpatient data in their model. While some models using only inpatient data performed slightly better [29], they lack the ability to make statements about the relevance of health care maintenance outside of the hospital setting to readmission. Other models predicting all-cause readmission using inpatient data were from the United States and considered 30-day admission instead of 1 year. Nonetheless, the predictive performance of our model for 1-year all-cause readmission was slightly better than these, with a C-statistic of 0.68, instead of 0.62 and 0.64 [30,31]. Our best model also outperformed an untargeted analysis in the same data [32], potentially demonstrating the performance gain that

can come with careful targeting of both population and model, though the relative contribution of each remains unclear.

Overall, many of the predictors for readmission that we identified as important have previously been mentioned by other studies. Surprisingly, in our data set, pantoprazole was the most important predictor for all-cause readmission. This variable was not mentioned in literature on predictors of readmission in patients with HF before. However, pantoprazole should be probably considered a proxy for overall disease severity. Proton pump inhibitors (PPIs) including pantoprazole are among the most commonly prescribed drugs in the German health care system [33]. PPIs are approved for short term (maximum 12 weeks) use to treat gastrointestinal acid–related disorders [34]. However, studies indicate that PPIs are overprescribed [35], and long-term use of PPIs is associated with increased risk for several adverse health outcomes such as fractures [36] and pneumonia [37]. Noncardiovascular comorbidities are strongly associated with readmission in HF, with pulmonary diseases and bone or joint disorders having the highest proportion among noncardiovascular causes for readmission [38]. Given these findings, exposure to pantoprazole may be a plausible predictor for 1-year all-cause readmission in patients with HF as seen in our data. Nevertheless, these results must be interpreted with caution and should be confirmed in future studies.

Being male was a risk factor for readmission, consistent both with some other HF readmission literature that uses a longer readmission period [39]. Age at which HF occurred was also an important predictor. In univariate and regression models, increasing age was associated with the risk of readmission, an effect that is potentially consistent with previous reported relationships between frailty and HF readmission [40], although this requires further study. We also reported the association of degree of rurality as an important predictor. While other studies have included variables such as distance to the nearest hospital [18], and both the association between rurality with health [41] and rurality with HF prevalence [42], we are, to our knowledge, the first to report this as a relevant predictor for HF readmission. One previous study found that socioeconomically deprived areas had no significant effect on 1-year all-cause readmission in patients with HF using logistic regression [43], but this study did not consider good geographical accessibility of a hospital. Other important predictors such as diabetes, COPD, and coronary disease have been widely and consistently reported in the literature [44-46].

Interestingly, enrolled in a DMP was associated with risk of 1-year readmission in our data. This conflicts with previously published data, also from the AOK Routine Data set Baden-Württemberg, that found that participation in a DMP for diabetes mellitus type 2 was protective in patients with HF against all-cause readmission over an 8-year period [47]. In our analysis, among those not readmitted within 1 year, the rates of participation in DMPs increased with time until readmission. Therefore, we posit that in the short term, participation in DMPs is a marker for chronic disease requiring care and therefore associated with readmission in some patients, but for those who are not quickly readmitted, DMPs can reduce the likelihood of readmission in long term. However, this needs to be confirmed in future studies.

Limitations

This study has several important limitations. First, we are unaware of any events that occur outside those stated in the data. While we do not expect significant numbers of HF admissions that are undocumented in the data, we cannot be sure whether any occur. Similarly, we have no control over the accuracy of the data set. While we attempted quality control steps to account for clearly impossible data, data points that fell within the plausible spectrum but were incorrect were not adjusted. In addition, due to the nature of health insurance data, no clinical information on HF severity was included. This means that we are able to distinguish the reliability of our predictions for an individual with early versus late stage HF. However, as shown by Desai et al [48], adding electronic health record information to prediction of HF readmission in ML models did not improve model performance. Another limitation is the lack of cardiovascular imaging and measurement. Due to the nature of insurance data, information types that may be relevant in predicting HF readmission, including echocardiography, electrocardiograms, and other imaging data were not available. While other studies have shown these may be relevant for predicting HF, their lack of availability in insurance data is expected. Nevertheless, we recognize that different subsets of patients with HF by ejection fraction may have different sets of predictors that we were unable to evaluate in this study. We also excluded individuals who had HF before 50 years or who lived in nursing facilities. Our conclusions therefore may not be relevant to these populations. One final limitation is the generalizability of our results to the whole German population. Although the AOK Baden-Württemberg covers nearly half of the population in Baden-Württemberg, it is not clear if similar patterns would be apparent within other SHIs or if the characteristics of patients who choose different SHIs would somehow affect this. It is also not clear whether these results are relevant to countries that lack SHIs.

Conclusions

This study shows that outpatient data from SHI can provide important information for the prediction of all-cause and HF-specific readmission after first admission for HF. It also highlights the relevance of social factors, DMPs, and concerns regularly addressed by primary care physicians in predicting readmission. Future prospective studies are needed to evaluate whether ML models of readmission are accurate in real time and relevant for clinical care.

Acknowledgments

The authors would like to acknowledge Jan D Lanzer for his insightful comments in discussions that shaped this project. This study was founded by the German Innovation Funds project PREMISE (grant 01VSF18019), and no authors received personal

funding. The funding body did not play any role in the design of the study, collection, analyses, and interpretation of data or the manuscript. RTL is funded by the Klaus Tschira Stiftung through the Informatics for Life Consortium. For the publication fee, we acknowledge financial support by Heidelberg University.

Data Availability

The data sets generated during and analyzed during this study are not publicly available, as they are proprietary of the health insurance company AOK Baden-Württemberg (third-party data), and we are legally not allowed to share these data. Permission to use the data set was granted by the AOK Baden-Württemberg for the specified purpose of readmission analyses within German Innovation Funds project PREMISE according to § 92a (2) Volume V of the Social Insurance Code (§ 92a Abs 2, SGB V—Fünftes Buch Sozialgesetzbuch; grant number 01VSF18019) [49]. Requests to use the data should be addressed to AOK Baden-Württemberg [50]. We hereby confirm that the authors had no special access to the data and that qualified researchers can request access to the data in the same way the authors obtained it.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Population demographics stratified by training or testing set and heart failure (HF)–specific readmission status, as well as information about HF-specific readmission models.
[DOCX File , 43 KB - cardio_v8i1e54994_app1.docx]

References

1. Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. Nat Rev Cardiol 2016;13(6):368-378 [FREE Full text] [doi: 10.1038/nrcardio.2016.25] [Medline: 26935038]
2. Ruff C, Gerharz A, Groll A, Stoll F, Wirbka L, Haefeli WE, et al. Disease-dependent variations in the timing and causes of readmissions in Germany: a claims data analysis for six different conditions. PLoS One 2021;16(4):e0250298 [FREE Full text] [doi: 10.1371/journal.pone.0250298] [Medline: 33901203]
3. Curtis LH, Greiner MA, Hammill BG, Kramer JM, Whellan DJ, Schulman KA, et al. Early and long-term outcomes of heart failure in elderly persons, 2001-2005. Arch Intern Med 2008 Dec 08;168(22):2481-2488 [FREE Full text] [doi: 10.1001/archinte.168.22.2481] [Medline: 19064833]
4. Liao L, Allen LA, Whellan DJ. Economic burden of heart failure in the elderly. Pharmacoeconomics 2008;26(6):447-462. [doi: 10.2165/00019053-200826060-00001] [Medline: 18489197]
5. Arundel C, Lam PH, Khosla R, Blackman MR, Fonarow GC, Morgan C, et al. Association of 30-day all-cause readmission with long-term outcomes in hospitalized older Medicare beneficiaries with heart failure. Am J Med 2016 Nov;129(11):1178-1184 [FREE Full text] [doi: 10.1016/j.amjmed.2016.06.018] [Medline: 27401949]
6. Saito M, Negishi K, Marwick TH. Meta-analysis of risks for short-term readmission in patients with heart failure. Am J Cardiol 2016;117(4):626-632. [doi: 10.1016/j.amjcard.2015.11.048] [Medline: 26772444]
7. Kreis K, Neubauer S, Klora M, Lange A, Zeidler J. Status and perspectives of claims data analyses in Germany-a systematic review. Health Policy 2016;120(2):213-226. [doi: 10.1016/j.healthpol.2016.01.007] [Medline: 26826756]
8. Statutory health insurance. GKV-Spitzenverband. URL: https://www.gkv-spitzenverband.de/english/statutory_health_insurance/statutory_health_insurance.jsp [accessed 2024-04-25]
9. Busse R, Blümel M, Knieps F, Bärnighausen T. Statutory health insurance in Germany: a health system shaped by 135 years of solidarity, self-governance, and competition. Lancet 2017;390(10097):882-897 [FREE Full text] [doi: 10.1016/S0140-6736(17)31280-1] [Medline: 28684025]
10. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. J Clin Epidemiol 2000;53(12):1258-1267. [doi: 10.1016/s0895-4356(00)00256-0] [Medline: 11146273]
11. Madelaire C, Gustafsson F, Kristensen SL, D'Souza M, Stevenson LW, Kober L, et al. Burden and causes of hospital admissions in heart failure during the last year of life. JACC Heart Fail 2019 Jul;7(7):561-570 [FREE Full text] [doi: 10.1016/j.jchf.2019.03.018] [Medline: 31248567]
12. Lee CS, Tkacs NC, Riegel B. The influence of heart failure self-care on health outcomes: hypothetical cardioprotective mechanisms. J Cardiovasc Nurs 2009;24(3):179-187; quiz 188 [FREE Full text] [doi: 10.1097/JCN.0b013e31819b5419] [Medline: 19279494]
13. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019;380(14):1347-1358. [doi: 10.1056/nejmra1814259]
14. Bazoukis G, Stavrakis S, Zhou J, Bollepalli SC, Tse G, Zhang Q, et al. Machine learning versus conventional clinical methods in guiding management of heart failure patients-a systematic review. Heart Fail Rev 2021 Jan;26(1):23-34 [FREE Full text] [doi: 10.1007/s10741-020-10007-3] [Medline: 32720083]

15. Yilmaz A, Hayıroğlu M, Salturk S, Pay L, Demircali AA, Coşkun C, et al. Machine learning approach on high risk treadmill exercise test to predict obstructive coronary artery disease by using P, QRS, and T waves' features. Curr Probl Cardiol 2023;48(2):101482. [doi: 10.1016/j.cpcardiol.2022.101482] [Medline: 36336117]
16. Hayıroğlu M, Altay S. The role of artificial intelligence in coronary artery disease and atrial fibrillation. Balkan Med J 2023;40(3):151-152 [FREE Full text] [doi: 10.4274/balkanmedj.galenos.2023.06042023] [Medline: 37025078]
17. Shin S, Austin PC, Ross HJ, Abdel-Qadir H, Freitas C, Tomlinson G, et al. Machine learning vs. conventional statistical models for predicting heart failure readmission and mortality. ESC Heart Fail 2021;8(1):106-115 [FREE Full text] [doi: 10.1002/ehf2.13073] [Medline: 33205591]
18. van der Galiën OP, Hoekstra RC, Gürgöze MT, Manintveld OC, van den Bunt MR, Veenman CJ, et al. Prediction of long-term hospitalisation and all-cause mortality in patients with chronic heart failure on Dutch claims data: a machine learning approach. BMC Med Inform Decis Mak 2021;21(1):303 [FREE Full text] [doi: 10.1186/s12911-021-01657-w] [Medline: 34724933]
19. Chamberlain AM, Dunlay SM, Gerber Y, Manemann SM, Jiang R, Weston SA, et al. Burden and timing of hospitalizations in heart failure: a community study. Mayo Clin Proc 2017;92(2):184-192 [FREE Full text] [doi: 10.1016/j.mayocp.2016.11.009] [Medline: 28160871]
20. Thünen Landatlas. URL: https://karten.landatlas.de/ [accessed 2023-05-08]
21. R Core Team. R: A Language and Environment for Statistical Computing. 2013. URL: http://www.R-project.org [accessed 2024-06-20]
22. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. J Open Source Softw 2019;4(43):1686 [FREE Full text] [doi: 10.21105/joss.01686]
23. Dowle M, Srinivasan A. data.table: Extension of ‘data.frame’. Data.table. 2023. URL: https://tinyurl.com/ymvzzjxm [accessed 2024-04-25]
24. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer Science & Business Media; 2009.
25. Lang M, Binder M, Richter J, Schratz P, Pfisterer F, Coors S, et al. mlr3: a modern object-oriented machine learning framework in R. J Open Source Softw 2019;4(44):1903 [FREE Full text] [doi: 10.21105/joss.01903]
26. Kuhn M. Building predictive models in R using the caret package. J Stat Soft 2008;28(5):1-26 [FREE Full text] [doi: 10.18637/jss.v028.i05]
27. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77 [FREE Full text] [doi: 10.1186/1471-2105-12-77] [Medline: 21414208]
28. Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - (Artikel 1 des Gesetzes v. 20. Dezember 1988, BGBl. I S. 2477) § 303e Datenverarbeitung. Bundesministerium der Justiz. URL: https://www.gesetze-im-internet.de/sgb_5/__303e.html [accessed 2024-06-20]
29. Golas SB, Shibahara T, Agboola S, Otaki H, Sato J, Nakae T, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. BMC Med Inform Decis Mak 2018;18(1):44 [FREE Full text] [doi: 10.1186/s12911-018-0620-z] [Medline: 29929496]
30. Allam A, Nagy M, Thoma G, Krauthammer M. Neural networks versus logistic regression for 30 days all-cause readmission prediction. Sci Rep 2019;9(1):9277 [FREE Full text] [doi: 10.1038/s41598-019-45685-z] [Medline: 31243311]
31. Frizzell JD, Liang L, Schulte PJ, Yancy CW, Heidenreich PA, Hernandez AF, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. JAMA Cardiol 2017;2(2):204-209 [FREE Full text] [doi: 10.1001/jamacardio.2016.3956] [Medline: 27784047]
32. Gerharz A, Ruff C, Wirbka L, Stoll F, Haefeli WE, Groll A, et al. Predicting hospital readmissions from health insurance claims data: a modeling study targeting potentially inappropriate prescribing. Methods Inf Med 2022;61(1-02):55-60 [FREE Full text] [doi: 10.1055/s-0042-1742671] [Medline: 35144291]
33. Schwabe U, Paffrath D, Ludwig WD, Klauber J. Arzneiverordnungs-Report 2018 [Drug Prescription Report 2018]. Berlin, Heidelberg: Springer-Verlag; 2018.
34. Katz PO, Dunbar KB, Schnoll-Sussman FH, Greer KB, Yadlapati R, Spechler SJ. ACG clinical guideline for the diagnosis and management of gastroesophageal reflux disease. Am J Gastroenterol 2022;117(1):27-56 [FREE Full text] [doi: 10.14309/ajg.0000000000001538] [Medline: 34807007]
35. Forgacs I, Loganayagam A. Overprescribing proton pump inhibitors. BMJ 2008;336(7634):2-3 [FREE Full text] [doi: 10.1136/bmj.39406.449456.BE] [Medline: 18174564]
36. Fraser LA, Leslie WD, Targownik LE, Papaioannou A, Adachi JD, CaMos Research Group. The effect of proton pump inhibitors on fracture risk: report from the Canadian Multicenter Osteoporosis Study. Osteoporos Int 2013;24(4):1161-1168 [FREE Full text] [doi: 10.1007/s00198-012-2112-9] [Medline: 22890365]
37. Ramsay EN, Pratt NL, Ryan P, Roughead EE. Proton pump inhibitors and the risk of pneumonia: a comparison of cohort and self-controlled case series designs. BMC Med Res Methodol 2013;13:82 [FREE Full text] [doi: 10.1186/1471-2288-13-82] [Medline: 23800078]

38. Dunlay SM, Redfield MM, Weston SA, Therneau TM, Long KH, Shah ND, et al. Hospitalizations after heart failure diagnosis a community perspective. J Am Coll Cardiol 2009;54(18):1695-1702 [FREE Full text] [doi: 10.1016/j.jacc.2009.08.019] [Medline: 19850209]
39. Hoang-Kim A, Parpia C, Freitas C, Austin PC, Ross HJ, Wijeysundera HC, et al. Readmission rates following heart failure: a scoping review of sex and gender based considerations. BMC Cardiovasc Disord 2020;20(1):223 [FREE Full text] [doi: 10.1186/s12872-020-01422-3] [Medline: 32408892]
40. Zheng PP, Yao SM, He W, Wan YH, Wang H, Yang JF. Frailty related all-cause mortality or hospital readmission among adults aged 65 and older with stage-B heart failure inpatients. BMC Geriatr 2021;21(1):125 [FREE Full text] [doi: 10.1186/s12877-021-02072-6] [Medline: 33593292]
41. van Dis J. Where We Live: Health Care in Rural vs Urban America. JAMA 2002;287(1):108. [Medline: 11754718]
42. Holstiege J, Akmatov MK, Störk S, Steffen A, Bätzing J. Higher prevalence of heart failure in rural regions: a population-based study covering 87% of German inhabitants. Clin Res Cardiol 2019;108(10):1102-1106 [FREE Full text] [doi: 10.1007/s00392-019-01444-8] [Medline: 30798347]
43. Al-Omary MS, Khan AA, Davies AJ, Fletcher PJ, Mcivor D, Bastian B, et al. Outcomes following heart failure hospitalization in a regional Australian setting between 2005 and 2014. ESC Heart Fail 2018;5(2):271-278 [FREE Full text] [doi: 10.1002/ehf2.12239] [Medline: 29265710]
44. Melbye H, Stylidis M, Solis JCA, Averina M, Schirmer H. Prediction of chronic heart failure and chronic obstructive pulmonary disease in a general population: the Tromsø study. ESC Heart Fail 2020;7(6):4139-4150 [FREE Full text] [doi: 10.1002/ehf2.13035] [Medline: 33025768]
45. Khan MS, Tahhan AS, Vaduganathan M, Greene SJ, Alrohaibani A, Anker SD, et al. Trends in prevalence of comorbidities in heart failure clinical trials. Eur J Heart Fail 2020;22(6):1032-1042 [FREE Full text] [doi: 10.1002/ejhf.1818] [Medline: 32293090]
46. Su A, Al'Aref SJ, Beecy AN, Min JK, Karas MG. Clinical and socioeconomic predictors of heart failure readmissions: a review of contemporary literature. Mayo Clin Proc 2019;94(7):1304-1320. [doi: 10.1016/j.mayocp.2019.01.017] [Medline: 31272573]
47. Sawicki OA, Mueller A, Klaaßen-Mielke R, Glushan A, Gerlach FM, Beyer M, et al. Strong and sustainable primary healthcare is associated with a lower risk of hospitalization in high risk patients. Sci Rep 2021;11(1):4349 [FREE Full text] [doi: 10.1038/s41598-021-83962-y] [Medline: 33623130]
48. Desai RJ, Wang SV, Vaduganathan M, Evers T, Schneeweiss S. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. JAMA Netw Open 2020;3(1):e1918962 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.18962] [Medline: 31922560]
49. Gemeinsamer Bundesausschuss-Innovationsfonds. URL: https://innovationsfonds.g-ba.de/ [accessed 2024-06-20]
50. AOK Baden-Württemberg. URL: https://www.aok.de/pk/bw/ [accessed 2024-06-20]

Abbreviations

ATC: Anatomical Therapeutic Chemical Classification
COPD: chronic obstructive pulmonary disease
DMP: disease management program
HF: heart failure
***ICD-10*:** *International Classification of Disease, 10th Revision*
ML: machine learning
PPI: proton pump inhibitor
RF: random forest
SHI: statutory health insurance

Edited by A Mavragani; submitted 29.11.23; peer-reviewed by M Hayıroğlu, I Lykhasenko, D Uppal; comments to author 29.02.24; revised version received 21.03.24; accepted 22.03.24; published 23.07.24.

Please cite as:
Levinson RT, Paul C, Meid AD, Schultz JH, Wild B
Identifying Predictors of Heart Failure Readmission in Patients From a Statutory Health Insurance Database: Retrospective Machine Learning Study
JMIR Cardio 2024;8:e54994
URL: https://cardio.jmir.org/2024/1/e54994
doi:10.2196/54994
PMID:

Original Paper

Formative Perceptions of a Digital Pill System to Measure Adherence to Heart Failure Pharmacotherapy: Mixed Methods Study

Peter R Chai[1,2,3,4], MD, MS; Jenson J Kaithamattam[1], BS; Michelle Chung[1], BA; Jeremiah J Tom[1], MBBS; Georgia R Goodman[1,4], BS; Mohammad Adrian Hasdianda[1], MD, MSc, MMSc; Tony Christopher Carnes[5], PhD; Muthiah Vaduganathan[6], MD, MPH; Benjamin M Scirica[6], MD, MPH; Jeffrey L Schnipper[7], MD

[1]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, United States

[2]Department of Psychosocial Oncology and Palliative Care, Dana Farber Cancer Institute, Boston, MA, United States

[3]The Koch Institute for Integrated Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, United States

[4]The Fenway Institute, Boston, MA, United States

[5]eTectRx, Gainesville, FL, United States

[6]Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States

[7]Division of Hospital Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States

Corresponding Author:
Peter R Chai, MD, MS
Department of Emergency Medicine
Brigham and Women's Hospital
75 Francis Street
Boston, MA, 02115
United States
Phone: 1 617 732 5640
Email: pchai@bwh.harvard.edu

Abstract

Background: Heart failure (HF) affects 6.2 million Americans and is a leading cause of hospitalization. The mainstay of the management of HF is adherence to pharmacotherapy. Despite the effectiveness of HF pharmacotherapy, effectiveness is closely linked to adherence. Measuring adherence to HF pharmacotherapy is difficult; most clinical measures use indirect strategies such as calculating pharmacy refill data or using self-report. While helpful in guiding treatment adjustments, indirect measures of adherence may miss the detection of suboptimal adherence and co-occurring structural barriers associated with nonadherence. Digital pill systems (DPSs), which use an ingestible radiofrequency emitter to directly measure medication ingestions in real-time, represent a strategy for measuring and responding to nonadherence in the context of HF pharmacotherapy. Previous work has demonstrated the feasibility of using DPSs to measure adherence in other chronic diseases, but this strategy has yet to be leveraged for individuals with HF.

Objective: We aim to explore through qualitative interviews the facilitators and barriers to using DPS technology to monitor pharmacotherapy adherence among patients with HF.

Methods: We conducted individual, semistructured qualitative interviews and quantitative assessments between April and August 2022. A total of 20 patients with HF who were admitted to the general medical or cardiology service at an urban quaternary care hospital participated in this study. Participants completed a qualitative interview exploring the overall acceptability of and willingness to use DPS technology for adherence monitoring and perceived barriers to DPS use. Quantitative assessments evaluated HF history, existing medication adherence strategies, and attitudes toward technology. We analyzed qualitative data using applied thematic analysis and NVivo software (QSR International).

Results: Most participants (12/20, 60%) in qualitative interviews reported a willingness to use the DPS to measure HF medication adherence. Overall, the DPS was viewed as useful for increasing accountability and reinforcing adherence behaviors. Perceived barriers included technological issues, a lack of need, additional costs, and privacy concerns. Most were open to sharing adherence data with providers to bolster clinical care and decision-making. Reminder messages following detected nonadherence were

perceived as a key feature, and customization was desired. Suggested improvements are primarily related to the design and usability of the Reader (a wearable device).

Conclusions: Overall, individuals with HF perceived the DPS to be an acceptable and useful tool for measuring medication adherence. Accurate, real-time ingestion data can guide adherence counseling to optimize adherence management and inform tailored behavioral interventions to support adherence among patients with HF.

(JMIR Cardio 2024;8:e48971) doi:10.2196/48971

KEYWORDS

behavioral interventions; cardiac treatment; digital pill system; heart failure medication; heart failure; ingestible sensors; medication adherence

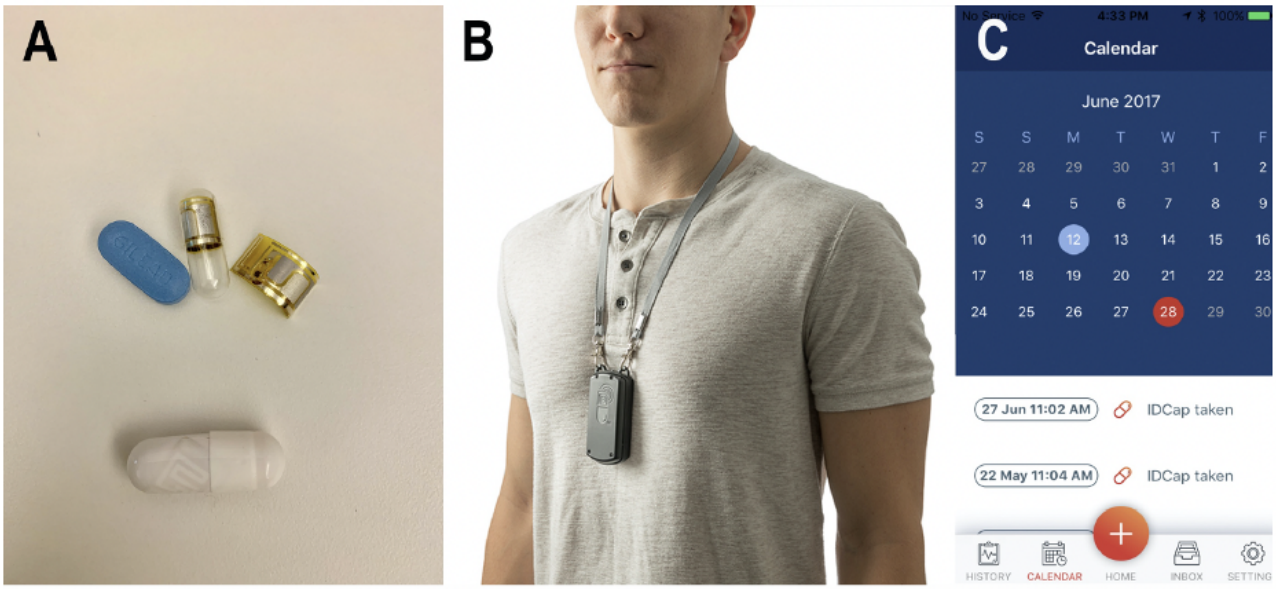
Introduction

Heart failure (HF) is one of the leading causes of morbidity and mortality in the United States, affecting approximately 6.2 million Americans [1,2]. In 2018, a total of 13.4% of deaths in the United States were attributed to HF [2]. HF is also one of the most common causes of hospitalization in individuals aged 65 years or older [3]. Among those admitted to the hospital, nearly one-fifth will be readmitted to the hospital for complications related to HF or other comorbidities within 30 days [4-6]. Pharmacologic management of HF focuses on increasing uptake and adherence to goal-directed quadruple medical therapy: angiotensin receptor-neprilysin inhibitor, β-blocker, mineralocorticoid receptor antagonist, and sodium-glucose co-transporter 2 inhibitors [7]. This strategy has demonstrated high efficacy for reducing hospital readmission and progression of HF and its associated cardiometabolic outcomes [8,9].

Medication nonadherence is a leading driver of worsening clinical outcomes in HF. Large longitudinal cohort studies have demonstrated that nonadherence to any pillar of HF pharmacotherapy is associated with increased all-cause mortality and an increased risk of 30-day hospital readmissions [10,11]. In a large, single-center, cross-sectional study, up to 15% of hospital readmissions in individuals with HF were associated with medication nonadherence [12]. Additionally, in individuals admitted to the hospital, 28% experience primary nonadherence to a component of HF pharmacotherapy as short as 1 week after discharge, with 24% experiencing persistent nonadherence at 30 days [13]. Given the close relationship between nonadherence and hospital readmission among individuals with HF, it is critical to continue to develop techniques that allow for the assessment of medication adherence in this population [14-20].

Current strategies for measuring adherence to HF pharmacotherapy include pharmacy refills, as measured by the medication possession ratio, and the number of subsequent days the patient has access to medications, as measured by the proportion of days covered [21,22]. This approach assesses overall adherence over periods of time, yet it is suboptimal in its capacity to capture daily challenges to adherence that may ultimately affect overall adherence and HF outcomes [23]. In contrast, one strategy for directly measuring daily adherence is a digital pill system (DPS; Figure 1). DPS technology is comprised of a gelatin capsule with an integrated radiofrequency emitter that overencapsulates the desired medication. Following ingestion of the digital pill, the radiofrequency emitter is activated by gastric chloride ions, which then projects a unique radio signal off the body that is acquired by a wearable device (Reader) [24,25]. The Reader stores and forwards ingestion data through low-energy Bluetooth to the user's smartphone and a clinician dashboard, enabling both patients and care teams to assess adherence patterns in real-time [26]. This strategy has been previously leveraged to measure oral pharmacotherapy adherence to antidiabetic and antihypertensive medications [27-30].

Figure 1. Components of the digital pill system (DPS; ID-Cap System; etectRx). (A) A radio frequency identification-tagged capsule with pill, (B) the Reader device worn on a lanyard over the neck, and (C) the smartphone app displaying the details of a digital pill ingestion.



To understand potential user responses to the DPS and inform future research involving this technology among individuals with HF, we conducted brief quantitative assessments and semistructured qualitative interviews to explore perceived facilitators of and barriers to the use of a DPS that measures adherence to HF pharmacotherapy.

Methods

Participants

All participants met the following inclusion criteria: (1) aged 18 years or older, (2) admitted to inpatient general medical or cardiology services with a diagnosis of HF, and (3) currently on oral HF pharmacotherapy. Individuals were excluded if they (1) had a history of heart transplant, (2) had an implanted left ventricular assist device, (3) were non-English speaking, or (4) were admitted to an intensive care unit.

Procedures

Participants were recruited in-person at a large, urban academic quaternary care hospital in Boston, Massachusetts, where patients with HF are admitted to either the general medical service or cardiology services; both inpatient teams independently manage patients with standardized treatment algorithms. Participants were not previously known to or in direct clinical care with any members of the study team. All study procedures were conducted in-person in a private area at the hospital while participants were inpatient.

Following verbal consent, participants completed a digitally recorded, semistructured qualitative interview with a bachelor's-level research assistant (either male or female) trained in qualitative interviewing techniques (JJK and MC). Interviews ranged from 23 minutes to 64 minutes in length (mean duration of 39 minutes). We adhered closely to the Consolidated Criteria for Reporting Qualitative Research (COREQ) guidelines (Multimedia Appendix 1) [31]. Study staff explained the components and functionality of the DPS (ID-Cap System; etectRx) in detail. Debrief documents were written after each interview and shared with the study team, who assessed for thematic saturation. Participants also completed a brief quantitative assessment. Following the completion of all study visit procedures, remuneration was provided. Study procedures were completed from April to August 2022.

Measures

Qualitative Interview

A qualitative interview guide (Multimedia Appendix 2) was developed by the study team members with expertise in the DPS, goal-directed medical therapy for HF, medication adherence, and technology development (PRC, JLS, MV, and BMS). Questions explored baseline adherence to HF medications and current adherence strategies, initial responses to DPS technology and messaging infrastructure through the DPS app or SMS text messaging, perceived facilitators of and barriers to DPS use, and perceptions of data privacy in the DPS context. Following an overview of the DPS technology and component parts, participants were asked whether they would be willing to use the DPS for HF adherence monitoring; this question was used to evaluate overall acceptance of the technology. The interview guide was piloted for completeness among members of the study team before implementation. Sample interview questions are provided in Table 1.

Table 1. Sample qualitative interview content areas, questions, and probes used during the study.

Content area	Sample probes
Current adherence strategies	• How long have you been prescribed a diuretic or SGLT2i[a]? • How have you tried to remember to take your medications? • What kind of barriers do you face to taking your medications on time?
DPS[b] technology	• What are your initial reactions to the digital pill? • Are there design factors to the digital pill and Reader that prevent you from wanting to use it? • Why would these factors prevent your use of digital pills?
DPS messaging components	• Tell me about situations you would like to receive notifications about your adherence. • What kind of messages would you want to receive in relation to the digital pill?
Data privacy and sharing	• The digital pill allows your provider or study team to view your adherence. What do you think of this? • What concerns do you have regarding the privacy of your adherence data? • Who do you think should have access to your adherence data? Why?
Acceptance of and willingness to use the DPS	• Given what you know, would you be willing to use the digital pill? Why or why not?

[a]SGLT2i: sodium-glucose co-transporter 2 inhibitors.

[b]DPS: digital pill system.

Quantitative Assessment

Quantitative assessments collected data surrounding sociodemographics and HF history. Participants were asked to estimate their adherence to HF medications over the past 3 months on a 0% to 100% sliding scale. We also provided a list of common medication adherence systems (eg, pill boxes, automated phone reminders, and smartphone apps) to assess previous use of such adherence strategies. These questionnaires were developed by the study team, which also supervised participants in completing the baseline assessment.

We used 3 subscales of the previously validated Media Technology Usage and Attitudes Scale (MTUAS) to measure attitudes toward technology: the positive attitudes subscale (6 items, eg, “With technology anything is possible”), the negative attitudes subscale (3 items, eg, “New technology makes life more complicated”), and the anxiety or dependence on technology subscale (3 items, eg, “I get anxious when I don’t have my cell phone”) [32]. Items were rated on a 5-point Likert scale (1=strongly disagree and 5=strongly agree). Score ranges were 1-5 for each subscale, with higher scores indicating more positive attitudes, more negative attitudes, and more technological anxiety and dependence [32]. The final quantitative assessment was cognitively tested among the study team to ensure clarity of questions before deployment with participants.

Analyses

Descriptive statistics were calculated to characterize the sample. Qualitative interviews were professionally transcribed and scrubbed of identifiers. Applied thematic analysis was used to code and analyze the interviews [33]. As part of the applied thematic analysis approach, 3 study team members (JJK, JJT, and GRG) reviewed all interview transcripts in order to iteratively generate a coding framework using a combination of the interview guide questions and data from the interviews themselves. Parent codes and subcodes were iteratively added to the coding framework throughout the transcript review process, and the final coding framework was then reviewed and revised by the study team before the formal coding of transcripts for the purpose of identifying qualitative domains and themes. Our 2 independent coders (JJK and JJT) double-coded 25% of the transcripts to establish interrater reliability; a κ score of >0.8 was used to establish adequate reliability between the coders, and this threshold was met. Study team members (JJK, JJT, and GRG) reviewed and compared coding throughout this process to discuss and resolve discrepancies, with oversight from the study’s principal investigator (PRC). Following the resolution of all coding discrepancies in double-coded transcripts, the coders (JJK and JJT) then independently coded the remaining 75% of transcripts. An audit trail of computerized coding was maintained. Salient quotes from the interviews were extracted, discussed with a subset of the study team (JJK, JJT, PRC, and GRG) to identify major domains and themes, and then disseminated to the entire study team for review. Coding was facilitated by NVivo software (QSR International).

Ethical Considerations

All study procedures were approved by the Mass General Brigham Institutional Review Board (2022P000545). We obtained written informed consent from all study participants. Study data were anonymized, and all study participants were only identified by a unique study identification number. Transcripts of interviews were scrubbed of any identifiers before analysis. Participants were compensated US $40 at the completion of interviews.

Results

Participant Characteristics

Over the study period, 96 individuals met the inclusion criteria. Of these, 43 (45%) were discharged before they could be

approached by the study team. Of the remaining 53 individuals, 12 (23%) were unavailable for consent, and 21 (40%) declined to participate. The reasons provided for declining participation included the time commitment for study procedures (n=2), general lack of interest (n=10), lack of knowledge of current medications (n=1), perception that they did not match the target study population (n=1), dissatisfaction with current clinical care (n=1), and reason unknown (n=6). A total of 20 participants consented and completed all study procedures (mean age 68, SD 14.3 years). The sample was predominantly female (n=11, 55%), White (n=13, 65%), and non-Hispanic (n=18, 90%). Full sociodemographic information is provided in Table 2.

Table 2. Sociodemographic characteristics of study participants (n=20).

Variable	Value
Age (years), mean (SD)	68 (14.3)
Sex, n (%)	
Male	9 (45)
Female	11 (55)
Race, n (%)	
Black or African American	6 (30)
White	13 (65)
Other	1 (5)
Ethnicity, n (%)	
Hispanic or Latino	2 (10)
Not Hispanic or Latino	18 (90)
Education, n (%)	
High school graduate or GED[a]	2 (10)
Some college	6 (30)
College degree	8 (40)
Some graduate school	2 (10)
Graduate or professional	2 (10)
Annual income (US $), n (%)	
6000-11,999	4 (20)
12,000-23,999	3 (15)
24,000-29,999	2 (10)
30,000-$59,999	3 (15)
≥60,000	8 (40)

[a]GED: general educational development.

Quantitative Results

Half (10/20, 50%) the sample had HF with preserved ejection fraction, and the other half (10/20, 50%) had HF with reduced ejection fraction. Most (11/20, 55%) were diagnosed with HF over 5 years ago, and half (10/20, 50%) had been admitted to the hospital multiple times due to HF in the past year. All (20/20, 100%) expressed at least some degree of concern regarding worsening HF. Self-reported adherence during the previous 3 months was high (mean 90.1%, SD 17.1%), and most (12/20, 60%) reported using a system to maintain adherence, with a standard pill box as the most common strategy (10/20, 50%). Finally, most participants (11/20, 55%) reported that visualizing their individual adherence patterns would motivate them to maintain adherence. Full HF status and pharmacotherapy adherence data are presented in Table 3.

Table 3. Heart failure (HF) status and pharmacotherapy adherence among study participants (n=20).

Variable	Value
HF status	
Duration of HF (years), n (%)	
<1	1 (5)
1-2	2 (10)
2-5	6 (30)
>5	11 (55)
Primary physician managing HF treatment, n (%)	
Cardiologist	13 (65)
Primary care physician	2 (10)
Does not know	4 (20)
Other	1 (5)
Number of prescribed HF medications, n (%)	
1	1 (5)
2-5	10 (50)
>5	9 (45)
Type of HF, n (%)	
HFpEF[a]	10 (50)
HFrEF[b]	10 (50)
Number of hospital admissions for HF over the last 12 months, n (%)	
0	2 (10)
1	7 (35)
2-5	10 (50)
>5	1 (5)
Number of physician encounters due to concerns surrounding worsening HF over last 12 months, n (%)	
0	7 (35)
1-5	10 (50)
6-10	1 (5)
11-20	2 (10)
Degree of concern about HF, n (%)	
Slightly concerned	2 (10)
Moderately concerned	5 (25)
Very concerned	4 (20)
Extremely concerned	9 (45)
Pharmacotherapy adherence	
Percentage of self-reported HF medication adherence over last 3 months, mean (SD)	90.1 (17.1)
Uses a system to maintain medication adherence, n (%)	
Yes	12 (60)
No	8 (40)
Medication adherence systems used, n (%)[c]	
Smart pill box	1 (8)
Pill organizer	10 (83)

Variable	Value
Smartphone-based reminders	3 (25)
Other	1 (8)
Visualization of adherence patterns would motivate medication adherence, n (%)	
Yes	151 (55)
No	7 (35)
Unsure	2 (10)
Willingness to use the DPS[d] , n (%)	
Yes	12 (60)
No	8 (40)

[a]HFpEF: heart failure with preserved ejection fraction.

[b]HFrEF: heart failure with reduced ejection fraction.

[c]Participants were provided with the opportunity to select multiple options, if applicable.

[d]DPS: digital pill system.

In terms of technology usage, three-quarters (15/20, 75%) of the sample owned a smartphone. MTUAS scores indicated positive attitudes toward technology (mean 4.2, SD 1.1) and a moderate degree of anxiety around being without technology and dependence on technology (mean 3.3, SD 1.4). Technology usage and MTUAS scores are provided in Table 4.

Table 4. Technology usage and the Media Technology Usage and Attitudes Scale (MTUAS) scores among study participants (n=20).

Variable	Value
Technology usage, n (%)	
Owns a smartphone	
Yes	15 (75)
No	5 (25)
Ever used a smartphone to communicate with medical care team	
Yes	13 (65)
No	7 (35)
Methods used to communicate with medical care team using a smartphone[a]	
Phone call	12 (92)
Through hospital portal (Patient Gateway)	10 (77)
Email	8 (62)
SMS text message	6 (46)
Other	2 (15)
MTUAS, mean (SD)	
Positive attitude toward technology subscale score	4.2 (1)
Negative attitude toward technology subscale score	3.0 (1)
Anxiety or dependence on technology subscale score	3.3 (1)

[a]Participants were provided with the opportunity to select multiple options.

Qualitative Results

Key findings surrounding the use of DPS technology for HF pharmacotherapy adherence emerged across the following major domains: (1) initial responses to the DPS, perceived barriers to use, and overall willingness to use the technology; (2) perceptions around privacy and sharing of DPS data; (3) responses to DPS messaging components; and (4) suggested improvements for future iterations. Multiple themes emerged within each domain; these are discussed in detail below.

Initial Responses, Perceived Barriers, and Overall Willingness to Use the DPS

Most participants perceived the DPS to be a novel, reliable tool for adherence measurement. They described the real-time data it generates as potentially useful for reinforcing adherence

behavior and noted that it would increase their sense of personal accountability for their HF regimen. Importantly, many participants described instances in which they were unsure whether they had taken their medications for the day and viewed the DPS as a valuable means for confirming past medication ingestions to avoid double dosing; this was interpreted as an indication of participants' perceived usefulness. After learning about the DPS, 60% (12/20) participants indicated a willingness to use the DPS to measure their HF pharmacotherapy adherence.

> *Absolutely I would use it. Because it's easier...It would help a whole lot. Because it would show [my physician] when or if I was adhering to the protocol. He'd know I'm taking my medicine or if I'm not.* [Aged 59 years, male]

> *I think it would be great—like a 30-day regime, make sure we're all on the same book kinda thing...If I was older, or I was gettin' blinky, or I didn't have caretakers or people looking out for me, it wouldn't be a bad idea.* [Aged 67 years, female]

> *I think it could be very useful for some people, and I doubt that I would use it right now in my present level of decline. But if I start having memory problems or if I ever start having problems taking medication, I'd be very interested in it.* [Aged 80 years, male]

Participants also identified a number of key barriers to DPS use. These included the perceived complexity of operating the technology, which was particularly salient among individuals who did not own smartphones. Some participants also described the Reader as large and potentially stigmatizing in the event that they needed to use the DPS in public. For some participants, the presence of electronics within the digital pill itself (ie, the radiofrequency emitter) raised questions around safety; however, most of these concerns were mitigated after participants were informed that the DPS in question (ID-Cap System; etectRx) had received Food and Drug Administration (FDA) clearance for use in humans. Other reported barriers to DPS uptake included potential costs associated with the device and a general lack of need for adherence support.

> *I don't have a comfort zone with technology. It scares me because I tried to learn, you know, particularly the phone, and I just get nervous...if I pick up something like that and do it, my mind just shuts down.* [Aged 80 years, male]

> *It does become a problem because you've probably already got everything else charged, and then you have to find a plug, figure out where you're gonna go with it. Then, if you got little kids, it's like, "What's that?" A whole bunch of headache.* [Aged 46 years, female]

Perceptions Around Privacy and Sharing of DPS Data

Some participants reported privacy-related concerns, including a fear of unwarranted tracking or interdiction of their adherence data and the potential for tampering with data, as additional barriers to DPS use. In particular, these participants expressed worries about whether swallowing a pill containing a radiofrequency emitter could transmit unwanted personal information to others related to their medications, adherence behavior, location tracking, and other physiological data.

> *You're gonna have to sit in front of me and explain to me how it's secure. What is making that radio frequency secure? Because I'm not just gonna randomly believe somebody that says, "Oh, well, you're gonna swallow this magic pill. It's gonna have a motherboard inside and it's gonna randomly broadcast to an outside device, and tell people what medications you're on, what you're taking, when you're taking it—and potentially additional information about it."* [Aged 58 years, female]

Despite expressing some concerns around data transmission and privacy, overall, participants expressed a desire for their DPS adherence data to be shared with their clinical care teams, given its importance for preventing the progression of HF. They reported that sharing DPS data with providers would be more reliable than self-reporting adherence and that it could be used to guide conversations around medication side effects, additional adherence or behavioral support that may be needed, and adjustments to medication regimens, including in the setting of worsening disease. Other participants shared more mixed opinions; while these individuals were willing to share adherence data with providers, they were unsure if doing so would meaningfully impact their ongoing HF treatment.

> *If a doctor looks at it and sees that you're not taking your medication, well, of course, something's gonna have to be done...There's nothing bad about the data going to the doctor...it's all positive. It certainly can't hurt.* [Aged 71 years, male]

> *It's so important to let your physician know you're actually taking that medication as prescribed. So if something is not working, then they know there's no question that this person was adhering to the prescribed treatment. And maybe this medication is not working for them. Maybe they need to increase it or get another one.* [Aged 62 years, female]

Responses to DPS Messaging Components

Participants were presented with an overview of three types of messages that can be programmed within the DPS: (1) confirmatory messages, sent after each ingestion to indicate successful detection; (2) reminder messages, sent before a prespecified dosing window; and (3) nonadherence reminder messages, sent after a dosing window if no ingestion had been detected.

Most participants accepted confirmatory messages following ingestions and viewed them as a useful feature for instances in which they were unsure if they had correctly operated the DPS. Importantly, because HF pharmacotherapy consists of multiple medication regimens, participants emphasized the need for confirmatory messages to specify the name of the medication ingested. They also expressed a desire to customize the timing of confirmatory messages; while some preferred a confirmation after each ingestion, others preferred less frequent messages, such as only at the end of the day or the week, as part of an adherence summary. Relatedly, participants suggested that the

frequency of messages could increase or decrease over time, based on DPS-detected patterns of adherence and nonadherence.

> *I found the [confirmatory messages] a little annoying. I get too many text messages, so where it would be helpful is if I'd forgotten to take the medicine, then if I got a reminder in a text message to take my medicine, that would be great. Once I've done it, I don't need the confirmation.* [Aged 80 years, male]

Overall, the majority of participants viewed reminder messages—and in particular, reminder messages that follow nonadherence detected by the DPS—as one of the most important features of the technology. Most reported that changes in routine and forgetfulness were common reasons for missed doses and noted that just-in-time reminders would be helpful for maximizing the potential for adherence in the moment, as well as for positive reinforcement around adherence behavior more generally. Some participants also noted that it would be useful to integrate their existing reminder systems, such as smartphone alarms, into DPS-based reminder messages in order to further reinforce adherence.

> *Usually what happens is, I don't know until the next day that I forgot to take [my medications], whereas, if I got a reminder at 9:00 p.m. saying, "Hey, you didn't take your nightly pills," that would be better, because then I could go take them.* [Aged 82 years, male]

> *So you don't need to beat somebody over the head, but they need to be told, "You missed your Lasix. This is a problem. You know, if you keep missing your Lasix, you could end up in the hospital." Like it needs to be made clear.* [Aged 58 years, female]

Participants also expressed an interest in customizing both the timing and content of reminder messages. In terms of timing, participants largely preferred a maximum of 2 messages proximal to each dosing window—for example, a reminder 30 minutes before the window and a follow-up reminder 15 minutes after the window if no ingestion was detected. Regarding the content of nonadherence reminders, participants reported an interest in simple messages indicating that they had forgotten to take their medication. Some also suggested that reminder messages could represent an opportunity to deliver HF-related educational information, especially related to the consequences of medication nonadherence.

> *You could do a snooze. You can pick, "Okay. Remind me five minutes before or five minutes after and during." I don't know. But let the person be able to choose how many reminders that they get.* [Aged 40 years, female]

> *If on the app, there's a little alarm that goes, "Hey, dummy, it's time to take your pill," and then I take a pill, and it monitors me taking the pill, then that's pretty much all you could ask.* [Aged 63 years, male]

Suggested Improvements for Future Iterations

Most recommendations focused on technological and design-based improvements to the Reader that would improve the user experience. Suggested enhancements included a new form factor that could integrate into typical clothing (eg, a pocket clip, wristband, smartphone case, or necklace). Participants also suggested that integrating additional features into the Reader, such as a voice assistant and colored lights to indicate adherence and reminders, would be helpful for individuals who do not carry a smartphone. Customization of the exterior casing of a Reader was also proposed, as was a stand-alone device that could provide adherence feedback independent of a smartphone. Finally, participants emphasized that future iterations of the system should come with detailed information around security protections and clear instructions for use.

> *I'd rather put it in my pocket or hold it in my hand. The best is just being able to plug it in and forget it...just because it's something you don't have to worry about anymore. I mean I have things plugged in around my house...I don't think anything about 'em—they're doing their job and that's all I have to do.* [Aged 82 years, male]

> *In the directions, I would want to be told that it's not harmful and why it's not harmful...I would like to know how long this is [for]. Like the directions say, if you take this gelatin pill...it will stay inside you for three months and it'll help us to track this for three months. I would like very clear instructions on how to use it.* [Aged 58 years, female]

Discussion

Overview

HF is one of the leading causes of hospital readmissions and mortality in the world [1,2]. One key pillar in efforts to optimize medical management of HF includes maximizing adherence to pharmacotherapy. While DPS technology has previously been shown to accurately measure medication adherence across a wide spectrum of diseases, its efficacy has yet to be described in the context of HF treatment [27-30]. This qualitative investigation provides formative data surrounding the acceptance and design of a DPS that directly measures HF pharmacotherapy adherence. Findings indicate that participants were accepting of the DPS overall and perceived the system as a tool for enhancing accountability and providing data to inform the ongoing medical management of HF. Personalized adherence reminders were identified as a key component of the system. These data demonstrate the potential for DPS deployment to measure adherence among individuals with HF.

Principal Findings

After learning about the DPS, 60% (12/20) of participants perceived the system as an "acceptable" strategy to measure HF pharmacotherapy adherence. Participants also expressed the "usefulness" of the DPS based on its perceived ability to motivate adherence, provide accountability, and avoid double-dosing. For most individuals, having incontrovertible evidence of their adherence (or nonadherence)—especially in the context of clinical care, where their DPS adherence data could aid discussions with their HF physicians and guide future medication decisions—was perceived as the most valuable benefit. These qualitative findings reinforce other proposed

benefits in the literature of leveraging real-time adherence data to not only address medication adherence in chronic disease but potentially enhance the patient-physician relationship by providing key data to ground conversations surrounding disease progression [34]. This concept was reflected in the quantitative portion of the study, where 55% (11/20) of individuals considered having a visual record of ingestion patterns over time as a motivating factor to continue to maintain medication adherence. These perceptions are consistent with other investigations that suggest individuals with other chronic diseases find value in DPS-based adherence data as a technique to guide pharmacotherapy [35-37]. Together, our data suggests that future research investigations should seek to understand the feasibility of real-world DPS operation among individuals with HF, as well as evaluate the impact of adherence metrics on disease progression and treatment regimens in both research and clinical deployment contexts.

Efforts to optimize the design of DPS technology to measure adherence to HF medication should also include a customizable messaging architecture that responds to detected patterns of adherence. Based on our data, messaging components should include confirmation messages to help individuals recognize that they correctly operated the DPS and recorded their medication ingestion. Most importantly, messaging modules should include nonadherence reminder messages that respond to adherence patterns from the DPS, which participants identified as a critical and valuable component of the system. A major theme that emerged from our interviews was participants' desire for control over both the timing and content of reminder messages related to their adherence patterns. While some wanted daily or even more frequent messaging that would coach them through adherence lapses, others preferred only on-demand access to their adherence data and less frequent feedback from the system. Importantly, some participants also reported that reminder messages could represent a potential method for providing educational information about HF and reinforcing DPS users' understanding of the consequences of medication nonadherence. These emerging themes demonstrate the importance of involving patients in the design and delivery of adherence interventions linked to digital health systems such as the DPS [29,38]. Barriers to the use of the DPS included discomfort with technology among some users and concerns about the privacy and security of their data.

Ultimately, participants viewed the DPS in its current iteration as usable, but they suggested several key improvements that would enable better integration into daily life. Some of these suggestions, including miniaturizing the Reader and providing alternative off-body systems that can collect adherence data, are currently under investigation in other ingestible sensor trials [39]. Additionally, participants expressed that DPS deployment should only occur alongside a detailed discussion with users about the safety and security of the system. While the DPS is FDA 510k cleared, participants emphasized the importance of providing users with data from past users of the system, particularly surrounding any DPS-related adverse events [25].

Limitations and Future Studies

This study had several limitations. First, the sample consisted of a small number of inpatients recruited as part of a convenience sample at a single hospital site. Qualitative data around patient experiences with HF and responses to DPS technology may vary across other health care institutions and patient populations. Second, perspectives from non–English-speaking individuals are missing, as this study only enrolled English-speaking participants; future investigations should explore responses to DPS technology in non-English speakers. Third, qualitative interviews explored perceptions of the technology among participants who did not ingest any digital pills or use the DPS themselves. The lived experiences of participants who use and operate the DPS in a clinical trial setting may differ.

Conclusions

In conclusion, this study demonstrates that individuals with HF perceived DPS technology to be an acceptable and useful tool for measuring medication adherence, informing our understanding of how this technology can be operationalized with this patient population in the real world. Importantly, this investigation also defined key boundary conditions for the physical design of the DPS as well as the structure of reminder messages that both support adherence and confirm the correct operation of the DPS. Finally, these formative data will help to inform best practices for future studies that develop interventions to support HF pharmacotherapy adherence and assess the efficacy of the DPS in this context.

Acknowledgments

The work on this study was funded by discretionary funds to PRC from the Department of Emergency Medicine, Brigham and Women's Hospital.

Data Availability

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

Authors' Contributions

PRC, JLS, MV, and BMS designed the study and developed the qualitative interview guide. JJK and MC conducted participant screening and approach, enrolment, and qualitative interviews. JJK, JJT, and GRG generated the codebook for evaluation of

results and facilitated the coding using NVivo. JJK, JJT, PRC, and GRG evaluated and identified major themes in the qualitative results. All authors read and approved the final manuscript.

Conflicts of Interest

PRC is funded by National Institutes of Health (NIH) K23DA044874 and DP2DA056107. PRC reports equity in Biobot Analytics and consulting fees from Syntis Bio. BMS reports institutional research grants to Brigham and Women's Hospital from Better Therapeutics, Boehringer Ingelheim, Merck, NovoNordisk, and Pfizer. Consulting fees from Abbvie (DSMB), AstraZeneca (DSMB), Boehringer Ingelheim, Better Therapeutics, Bristol Myers-Squibb, Elsevier Practice Update Cardiology, Esperion, Hanmi (DSMB), Lexeo (DSMB), Lexicon, NovoNordisk, and equity in Health at Scale and Doximity. MV reports grants from American Regent, Amgen, AstraZeneca, Roche Diagnostics, Galmed, Novartis, Bayer AG, Occlutech, and Impulse Dynamics. Consulting fees from American Regent, Amgen, AstraZeneca, Bayer AG, Baxter Healthcare, Boehringer Ingelheim, Chiesi, Cytokinetics, Lexicon Pharmaceuticals, Novartis, Merck, NovoNordisk, Pharmacosmos, Relypsa, Roche Diagnostics, Sanofi, and Tricog Health. Payment for lectures, presentations, speakers, and manuscript writing comes from AstraZeneca, Boehringer Ingelheim, Cytokinetics, Novartis, and Roche Diagnostics. JLS received funding from Synapse Medicine to conduct an investigator-initiated study of their medication decision support software and a stipend from the American Society of Health-System Pharmacists to develop a web-based course on medication history taking.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 417 KB - cardio_v8i1e48971_app1.pdf]

Multimedia Appendix 2
Qualitative interview guide.
[DOCX File , 27 KB - cardio_v8i1e48971_app2.docx]

References

1. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, et al. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. Circ Heart Fail 2013;6(3):606-619 [FREE Full text] [doi: 10.1161/HHF.0b013e318291329a] [Medline: 23616602]
2. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics-2020 update: a report from the American Heart Association. Circulation 2020;141(9):E139-E596 [FREE Full text] [doi: 10.1161/cir.0000000000000746] [Medline: 31992061]
3. Ambrosy AP, Fonarow GC, Butler J, Chioncel O, Greene SJ, Vaduganathan M, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. J Am Coll Cardiol 2014;63(12):1123-1133 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.053] [Medline: 24491689]
4. Healthcare Cost and Utilization Project (HCUPnet). Agency for Healthcare Research and Quality. URL: https://datatools.ahrq.gov/hcupnet/ [accessed 2020-08-05]
5. Donzé J, Lipsitz P, Bates DW, Schnipper JL. Causes and patterns of readmissions in patients with common comorbidities: retrospective cohort study. BMJ 2013;347:f7171 [FREE Full text] [doi: 10.1136/bmj.f7171] [Medline: 24342737]
6. Desai NR, Ross JS, Kwon JY, Herrin J, Dharmarajan K, Bernheim SM, et al. Association between hospital penalty status under the hospital readmission reduction program and readmission rates for target and nontarget conditions. JAMA 2016;316(24):2647-2656 [FREE Full text] [doi: 10.1001/jama.2016.18533] [Medline: 28027367]
7. Greene SJ, Butler J, Metra M. Another reason to embrace quadruple medical therapy for heart failure: medications enabling tolerance of each other. Eur J Heart Fail 2021;23(9):1525-1528 [FREE Full text] [doi: 10.1002/ejhf.2301] [Medline: 34263507]
8. Docherty KF, Bayes-Genis A, Butler J, Coats AJS, Drazner MH, Joyce E, et al. The four pillars of HFrEF therapy: is it time to treat heart failure regardless of ejection fraction? Eur Heart J Suppl 2022;24(Suppl L):L10-L19 [FREE Full text] [doi: 10.1093/eurheartjsupp/suac113] [Medline: 36545228]
9. Sharma A, Verma S, Bhatt DL, Connelly KA, Swiggum E, Vaduganathan M, et al. Optimizing foundational therapies in patients with HFrEF: how do we translate these findings into clinical care? JACC Basic Transl Sci 2022;7(5):504-517 [FREE Full text] [doi: 10.1016/j.jacbts.2021.10.018] [Medline: 35663626]
10. Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, et al. Factors identified as precipitating hospital admissions for heart failure and clinical outcomes: findings from OPTIMIZE-HF. Arch Intern Med 2008;168(8):847-854 [FREE Full text] [doi: 10.1001/archinte.168.8.847] [Medline: 18443260]
11. Fitzgerald AA, 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;17(8):664-669. [doi: 10.1016/j.cardfail.2011.04.011] [Medline: 21807328]

12. Chin MH, Goldman L. Factors contributing to the hospitalization of patients with congestive heart failure. Am J Public Health 1997;87(4):643-648 [FREE Full text] [doi: 10.2105/ajph.87.4.643] [Medline: 9146445]
13. Fallis BA, Dhalla IA, Klemensberg J, Bell CM. Primary medication non-adherence after discharge from a general internal medicine service. PLoS One 2013;8(5):e61735 [FREE Full text] [doi: 10.1371/journal.pone.0061735] [Medline: 23658698]
14. Packer M, Butler J, Zannad F, Filippatos G, Ferreira JP, Pocock SJ, et al. Effect of empagliflozin on worsening heart failure events in patients with heart failure and preserved ejection fraction: EMPEROR-preserved trial. Circulation 2021;144(16):1284-1294 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.121.056824] [Medline: 34459213]
15. McMurray JJV, DeMets DL, Inzucchi SE, Køber L, Kosiborod MN, Langkilde AM, et al. A trial to evaluate the effect of the sodium-glucose co-transporter 2 inhibitor dapagliflozin on morbidity and mortality in patients with heart failure and reduced left ventricular ejection fraction (DAPA-HF). Eur J Heart Fail 2019;21(5):665-675 [FREE Full text] [doi: 10.1002/ejhf.1432] [Medline: 30895697]
16. Omar M, Jensen J, Ali M, Frederiksen PH, Kistorp C, Videbæk L, et al. Associations of empagliflozin with left ventricular volumes, mass, and function in patients with heart failure and reduced ejection fraction: a substudy of the empire HF randomized clinical trial. JAMA Cardiol 2021;6(7):836-840 [FREE Full text] [doi: 10.1001/jamacardio.2020.6827] [Medline: 33404637]
17. Butler J, Usman MS, Khan MS, Greene SJ, Friede T, Vaduganathan M, et al. Efficacy and safety of SGLT2 inhibitors in heart failure: systematic review and meta-analysis. ESC Heart Fail 2020;7(6):3298-3309 [FREE Full text] [doi: 10.1002/ehf2.13169] [Medline: 33586910]
18. Ruppar TM, Cooper PS, Mehr DR, Delgado JM, Dunbar-Jacob JM. Medication adherence interventions improve heart failure mortality and readmission rates: systematic review and meta-analysis of controlled trials. J Am Heart Assoc 2016;5(6):e002606 [FREE Full text] [doi: 10.1161/JAHA.115.002606] [Medline: 27317347]
19. Ofori-Asenso R, Sahle BW, Chin KL, Mazidi M, Ademi Z, de Bruin ML, et al. Poor adherence and persistence to sodium glucose co-transporter 2 inhibitors in real-world settings: evidence from a systematic review and meta-analysis. Diabetes Metab Res Rev 2021;37(1):e3350. [doi: 10.1002/dmrr.3350] [Medline: 32447808]
20. Ofori-Asenso R, Liew D, Lalic S, Mazidi M, Magliano DJ, Ademi Z, et al. Adherence, persistence, and switching among people prescribed sodium glucose co-transporter 2 inhibitors: a nationwide retrospective cohort study. Adv Ther 2019;36(11):3265-3278. [doi: 10.1007/s12325-019-01077-3] [Medline: 31482509]
21. Qin X, Hung J, Knuiman MW, Briffa TG, Teng THK, Sanfilippo FM. Comparison of medication adherence measures derived from linked administrative data and associations with mortality using restricted cubic splines in heart failure patients. Pharmacoepidemiol Drug Saf 2020;29(2):208-218. [doi: 10.1002/pds.4939] [Medline: 31958191]
22. Sattler ELP, Lee JS, Perri M. Medication (re)fill adherence measures derived from pharmacy claims data in older Americans: a review of the literature. Drugs Aging 2013;30(6):383-399. [doi: 10.1007/s40266-013-0074-z] [Medline: 23553512]
23. Spreafico M, Gasperoni F, Barbati G, Ieva F, Scagnetto A, Zanier L, et al. Adherence to disease-modifying therapy in patients hospitalized for HF: findings from a community-based study. Am J Cardiovasc Drugs 2020;20(2):179-190. [doi: 10.1007/s40256-019-00367-z] [Medline: 31444666]
24. Chai PR, Goodman G, Bustamante MJ, Mohamed Y, Castillo-Mancilla J, Boyer EW, et al. Long-term stability of the electronic sensor component of a digital pill system in real-world storage settings. J Pharm Technol 2021;37(3):135-139 [FREE Full text] [doi: 10.1177/8755122520985219] [Medline: 34752557]
25. 510(k) premarket notification for etectRx. US Food and Drug Administration. 2019. URL: https://www.accessdata.fda.gov/Cdrh_docs/Pdf18/K183052.Pdf [accessed 2020-09-09]
26. Chai PR, Castillo-Mancilla J, Buffkin E, Darling C, Rosen RK, Horvath KJ, et al. Utilizing an ingestible biosensor to assess real-time medication adherence. J Med Toxicol 2015;11(4):439-444 [FREE Full text] [doi: 10.1007/s13181-015-0494-8] [Medline: 26245878]
27. Goodman GR, Vaz C, Albrechta H, Boyer EW, Mayer KH, O'Cleirigh C, et al. Ingestible electronic sensors for monitoring real-time adherence to HIV pre-exposure prophylaxis and antiretroviral therapy. Curr HIV/AIDS Rep 2022;19(5):433-445 [FREE Full text] [doi: 10.1007/s11904-022-00625-x] [Medline: 36048309]
28. Chai PR, Mohamed Y, Goodman G, Bustamante MJ, Sullivan MC, Najarro J, et al. Development of a digital pill and respondent behavioral intervention (PrEPSteps) for HIV pre-exposure prophylaxis adherence among stimulant using men who have sex with men. Transl Behav Med 2022;12(1):ibab117 [FREE Full text] [doi: 10.1093/tbm/ibab117] [Medline: 34453536]
29. Chai PR, Goodman GR, Bronzi O, Gonzales G, Baez A, Bustamante MJ, et al. Real-world user experiences with a digital pill system to measure PrEP adherence: perspectives from MSM with substance use. AIDS Behav 2022;26(7):2459-2468 [FREE Full text] [doi: 10.1007/s10461-022-03594-9] [Medline: 35089449]
30. Chai PR, Carreiro S, Innes BJ, Rosen RK, O'Cleirigh C, Mayer KH, et al. Digital pills to measure opioid ingestion patterns in emergency department patients with acute fracture pain: a pilot study. J Med Internet Res 2017;19(1):e19 [FREE Full text] [doi: 10.2196/jmir.7050] [Medline: 28087496]
31. 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;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

32. Rosen LD, Whaling K, Carrier LM, Cheever NA, Rokkum J. The media and technology usage and attitudes scale: an empirical investigation. Comput Human Behav 2013;29(6):2501-2511 [FREE Full text] [doi: 10.1016/j.chb.2013.06.006] [Medline: 25722534]
33. Guest G, MacQueen KM, Namey EE. Applied Thematic Analysis. Thousand Oaks, CA: Sage Publications; 2012.
34. Richey AG, Kovacs I, Browne S. Use of an ingestible, sensor-based digital adherence system to strengthen the therapeutic relationship in serious mental illness. JMIR Ment Health 2022;9(12):e39047 [FREE Full text] [doi: 10.2196/39047] [Medline: 36459392]
35. Chai PR, Mohamed Y, Bustamante MJ, Goodman GR, Najarro J, Castillo-Mancilla J, et al. DigiPrEP: a pilot trial to evaluate the feasibility, acceptability, and accuracy of a digital pill system to measure PrEP adherence in men who have sex with men who use substances. J Acquir Immune Defic Syndr 2022;89(2):e5-e15 [FREE Full text] [doi: 10.1097/QAI.0000000000002854] [Medline: 34753871]
36. Liu H, Wang Y, Huang Y, Xiong D, Shen J, Siqueiros L, et al. Ingestible sensor system for measuring, monitoring and enhancing adherence to antiretroviral therapy: an open-label, usual care-controlled, randomised trial. EBioMedicine 2022;86:104330 [FREE Full text] [doi: 10.1016/j.ebiom.2022.104330] [Medline: 36372731]
37. Browne SH, Behzadi Y, Littlewort G. Let visuals tell the story: medication adherence in patients with type II diabetes captured by a novel ingestion sensor platform. JMIR Mhealth Uhealth 2015;3(4):e108 [FREE Full text] [doi: 10.2196/mhealth.4292] [Medline: 26721413]
38. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015;22(6):754-756 [FREE Full text] [doi: 10.1111/acem.12692] [Medline: 25997375]
39. Next generation ingestible sensors for medication adherence measurement (NextGen). ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/NCT05592613 [accessed 2023-04-21]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DPS: digital pill system
FDA: Food and Drug Administration
HF: heart failure
MTUAS: Media Technology Usage and Attitudes Scale

Edited by A Mavragani; submitted 13.05.23; peer-reviewed by C Eaton, S Browne; comments to author 14.09.23; revised version received 19.10.23; accepted 22.12.23; published 15.02.24.

Please cite as:
Chai PR, Kaithamattam JJ, Chung M, Tom JJ, Goodman GR, Hasdianda MA, Carnes TC, Vaduganathan M, Scirica BM, Schnipper JL
Formative Perceptions of a Digital Pill System to Measure Adherence to Heart Failure Pharmacotherapy: Mixed Methods Study
JMIR Cardio 2024;8:e48971
URL: https://cardio.jmir.org/2024/1/e48971
doi:10.2196/48971
PMID:38358783

Original Paper

Association of Arterial Stiffness With Mid- to Long-Term Home Blood Pressure Variability in the Electronic Framingham Heart Study: Cohort Study

Xuzhi Wang[1], MS; Yuankai Zhang[1], MA; Chathurangi H Pathiravasan[1], PhD; Nene C Ukonu[1], MA; Jian Rong[2], PhD; Emelia J Benjamin[2,3,4], MD, ScM; David D McManus[5,6], MD, ScM; Martin G Larson[1,2], SD; Ramachandran S Vasan[2,7], MD; Naomi M Hamburg[8], MD; Joanne M Murabito[2,9], MD, ScM; Chunyu Liu[1*], PhD; Gary F Mitchell[10*], MD

[1]Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States

[2]Boston University's and National Heart, Lung, and Blood Institute's Framingham Heart Study, Boston, MA, United States

[3]Section of Cardiovascular Medicine, Department of Medicine, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States

[4]Department of Epidemiology, Boston University School of Public Health, Boston, MA, United States

[5]Cardiology Division, Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States

[6]Department of Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, United States

[7]Section of Preventive Medicine and Epidemiology, Department of Medicine, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States

[8]Section of Vascular Biology, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States

[9]Section of General Internal Medicine, Department of Medicine, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, United States

[10]Cardiovascular Engineering, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Chunyu Liu, PhD
Department of Biostatistics
Boston University School of Public Health
715 Albany Street
Boston, MA, 02118
United States
Phone: 1 6176385104
Email: liuc@bu.edu

Abstract

Background: Short-term blood pressure variability (BPV) is associated with arterial stiffness in patients with hypertension. Few studies have examined associations between arterial stiffness and digital home BPV over a mid- to long-term time span, irrespective of underlying hypertension.

Objective: This study aims to investigate if arterial stiffness traits were associated with subsequent mid- to long-term home BPV in the electronic Framingham Heart Study (eFHS). We hypothesized that higher arterial stiffness was associated with higher home BPV over up to 1-year follow-up.

Methods: At a Framingham Heart Study research examination (2016-2019), participants underwent arterial tonometry to acquire measures of arterial stiffness (carotid-femoral pulse wave velocity [CFPWV]; forward pressure wave amplitude [FWA]) and wave reflection (reflection coefficient [RC]). Participants who agreed to enroll in eFHS were provided with a digital blood pressure (BP) cuff to measure home BP weekly over up to 1-year follow-up. Participants with less than 3 weeks of BP readings were excluded. Linear regression models were used to examine associations of arterial measures with average real variability (ARV) of week-to-week home systolic (SBP) and diastolic (DBP) BP adjusting for important covariates. We obtained ARV as an average of the absolute differences of consecutive home BP measurements. ARV considers not only the dispersion of the BP readings around the mean but also the order of BP readings. In addition, ARV is more sensitive to measurement-to-measurement BPV compared with traditional BPV measures.

Results: Among 857 eFHS participants (mean age 54, SD 9 years; 508/857, 59% women; mean SBP/DBP 119/76 mm Hg; 405/857, 47% hypertension), 1 SD increment in FWA was associated with 0.16 (95% CI 0.09-0.23) SD increments in ARV of home SBP and 0.08 (95% CI 0.01-0.15) SD increments in ARV of home DBP; 1 SD increment in RC was associated with 0.14 (95% CI 0.07-0.22) SD increments in ARV of home SBP and 0.11 (95% CI 0.04-0.19) SD increments in ARV of home DBP. After adjusting for important covariates, there was no significant association between CFPWV and ARV of home SBP, and similarly, no significant association existed between CFPWV and ARV of home DBP (P>.05).

Conclusions: In eFHS, higher FWA and RC were associated with higher mid- to long-term ARV of week-to-week home SBP and DBP over 1-year follow-up in individuals across the BP spectrum. Our findings suggest that higher aortic stiffness and wave reflection are associated with higher week-to-week variation of BP in a home-based setting over a mid- to long-term time span.

(JMIR Cardio 2024;8:e54801) doi:10.2196/54801

KEYWORDS

arterial stiffness; mobile health; mHealth; blood pressure; blood pressure variability; risk factors

Introduction

Nearly half of US adults have hypertension [1]. The 2017 American College of Cardiology/American Heart Association blood pressure (BP) guidelines recommended out-of-office self-monitoring with home BP measurements to assist with hypertension diagnosis and management. Moreover, home-based BP measurements are stronger predictors of cardiovascular risk than office-based measurements [2]. BP fluctuates in response to everyday activities including exercise, mental stress, sleep, and other environmental stimuli. Parati et al [3] defined several types of blood pressure variability (BPV), including short-term BPV (eg, ambulatory BP monitoring within 24 hours), midterm BPV (eg, day-to-day BP monitoring in at least 3 days), and long-term BPV (eg, visit-to-visit BP monitoring over weeks to years). Elevated short-term BPV from ambulatory BP monitoring is associated with a higher risk of cardiovascular outcomes and all-cause mortality [4]. Day-to-day home BPV over 1 [5,6] to 4 weeks [7,8] is associated with cardiovascular risk and may identify persons at risk for cognitive decline [9]. Long-term BPV is associated with adverse cardiovascular events and mortality even after accounting for mean BP [10,11] in persons with and without hypertension [12]. However, week-to-week home BPV that is measured over the course of up to 1 year has not been well defined in the literature. In our study, we define mid- to long-term home BPV as the week-to-week home BPV collected during up to 1-year follow-up. Limited data are available on mid- to long-term home BPV in association with cardiovascular risk.

Arterial stiffness may be 1 biological mechanism linking BPV to cardiovascular disease risk [13]. Higher arterial stiffness is associated with a higher risk for incident hypertension [14] and is associated with both short-term and long-term adverse health outcomes including coronary disease events and heart failure [15,16]. Increased long-term visit-to-visit systolic BPV may contribute to the progression of arterial stiffness, regardless of mean BP levels [17]. Short-term (24-hour) BPV is associated with arterial stiffness [18]. Studies of mid- to long-term home BPV and arterial stiffness are limited; in one study, home BPV was correlated with a measure of arterial wave reflection in persons with high normal BP or hypertension [19]. Evaluation of the association of direct measures of arterial stiffness and wave reflection with home BPV over a mid- to long-term span is needed [3].

BPV indices have been proposed to calculate the overall variability as well as specific BP patterns [3], including SD, coefficient of variation (CV), average real variability (ARV), and variability independent of the mean. In our study, we obtained the ARV of week-to-week home BP from participants in the electronic Framingham Heart Study (eFHS) who returned digital BP data over up to 1-year follow-up. We investigated the association between arterial stiffness traits and the mid- to long-term home BPV, that is, the ARV of week-to-week home systolic blood pressure (SBP) and diastolic blood pressure (DBP) over up to 1-year follow-up from the eFHS participants. We hypothesize that higher arterial stiffness is associated with higher mid- to long-term home BPV over up to 1-year follow-up.

Methods

Study Participants

The Framingham Heart Study (FHS) is a multigenerational cohort study that began in 1948 with the original cohort enrolling 5209 residents from Framingham, Massachusetts. In 1971, the offspring cohort enrolled 5214 participants who were offspring of the original cohort and the spouses of these offspring. The FHS enrolled the Third Generation (Gen 3) cohort (N=4095), which included the grandchildren of original cohort from 2002 to 2005. During the same time, the FHS recruited and enrolled the Omni 2 cohort comprised of 410 multiethnic participants, and the New Offspring Spouse cohort (n=103) comprised of previously unenrolled parents of the Gen 3 participants. The participants in the Gen 3, Omni 2, and New Offspring Spouse cohorts underwent research exams approximately every 6 to 8 years. At examination 3 (2016 to 2019) participants underwent arterial tonometry testing and participants who owned a smartphone (including iPhone 4S or higher with iOS version 8.2 or higher or an Android phone beginning October 30, 2017) were invited to enroll in the eFHS. The eFHS participants downloaded a smartphone app and for participants using an iPhone, a Nokia Withings digital BP cuff was provided for home BP monitoring (Multimedia Appendix 1). The Nokia Withings digital cuff has been cleared for marketing by the Food and Drug Administration and it has been validated [20,21]. The eFHS participants downloaded the eFHS

smartphone app from the Apple Store with in-person help from the eFHS–trained staff or with written instructions provided by eFHS staff.

Participants with arterial tonometry measures who returned valid digital home BP readings as part of eFHS were eligible for inclusion in the study sample. A total of 3451 participants underwent arterial tonometry at examination 3. Of the 3451 participants, 2125 participated in the eFHS study. Among the eFHS participants, 1156 participants provided BP data using the digital BP cuff and the study smartphone app. We further excluded 299 participants for the following reasons: participants did not return BP measurements within the first 12 months of attending examination 3 (n=126), participants returned BP readings for <3 weeks (n=153), or participants had missing values in tonometry measures or covariates (n=20). After exclusion, 857 participants remained for subsequent statistical analysis (Multimedia Appendix 2). We further compared our final study sample with those who did not enroll in eFHS and those ineligible for inclusion in our final sample.

Assessment of Arterial Tonometry Measures

We examined 3 measures of arterial stiffness and wave reflection obtained from arterial tonometry including carotid-femoral pulse wave velocity (CFPWV), central forward pressure wave amplitude (FWA), and reflection coefficient (RC). Trained sonographers obtained the tonometry measurements using a standard protocol previously reported [22,23]. We chose to investigate CFPWV as this measure is the standard noninvasive measure of arterial stiffness recommended for vascular research [24]. CFPWV was calculated from carotid-femoral transit time delay and carotid-femoral transit distance adjusted for parallel transmission in the brachiocephalic and carotid arteries and aortic arch [23]. Wave separation was performed in the time domain [23]. FWA was defined as the amplitude of the forward pressure wave. RC was defined as the ratio of backward and forward wave amplitudes. As compared with CFPWV, FWA is comparably sensitive to aortic wall stiffness but is more sensitive to aortic diameter [14]. RC was assessed as a measure of relative wave reflection as described previously [25]. The magnitude of the RC relies on the degree of impedance mismatch between proximal and distal vessels. Impedance matching reduces RC and hence the amount of wave reflection at the interface between the aorta and branch vessels, which may result in the transmission of excessive pulsatile energy into the microcirculation where it can cause damage [25].

Assessment of Week-to-Week Home BP Measurement and BP Variability

eFHS participants were asked to measure and transmit home BP readings once each week using the Withings-Nokia digital BP cuff for up to 1 year following enrollment. Multimedia Appendix 1 displays the timeline of data collection for arterial tonometry measures, as well as for weekly home BP readings. eFHS staff demonstrated proper use of the digital BP cuff while the participant was in the Research Center. Written instructions were also provided as some participants chose to set up the digital cuff at home. Participants were asked to take a BP reading at about the same time of day and the same day of the week each week. Participants were advised to sit in a comfortable position and rest for 5 minutes before each BP reading was measured. Participants were instructed to avoid taking BP readings after doing exercise or consuming caffeinated beverages. All BP recordings were date- and time-stamped. We conducted quality control procedures in the following way: BP readings taken at the research center during training by eFHS staff, duplicate observations with identical BP readings, observations with values that likely represented spurious results including observations with DBP>SBP, SBP>250, DBP>140, SBP<70, or DBP<40 were excluded. To reduce the bias of week-to-week home BPV, we included eFHS participants who had BP readings for at least 3 weeks [26]. Because we aimed to investigate the association of antecedent arterial stiffness with week-to-week home BPV over a mid- to long-term time span, we excluded BP measurements assessed more than 1 year after enrollment.

Several common variability measures have been used to assess BPV, including SD and CV. Compared with these BPV indices, ARV not only considers the dispersion of the BP time series around the mean but also accounts for the order of BP readings [27,28]. Furthermore, SD and CV are sensitive to long-range variation, such as a progressive increase or decrease in BP, while being less sensitive to measurement-to-measurement variability. ARV was first proposed to calculate the 24-hour ambulatory BPV. In our study, we proposed to apply the ARV to week-to-week home SBP and DBP measurements over a mid- to long-term time span defined as up to 1 year. ARV was calculated as the average of absolute differences between the adjacent BP readings using the following formula:

where N is the number of weekly BP measurements and K is the order of weekly BP measurements.

Covariates

Clinical and laboratory variables were collected during examination 3 at the research center. BMI was calculated by dividing body weight (kg) by height (meters) squared. The current smoking variable was defined as self-reported smoking 1 or more cigarettes per day on average in the year before the examination. Lipid-lowering treatment variable was defined as a self-report of receiving lipid treatment in the past month before the examination. Antihypertensive medication variable was defined as a self-report of taking antihypertensive medication in the past month before the examination. Diabetes was defined as fasting plasma glucose ≥126 mg/dL or self-reported use of medications for diabetes. SBP and DBP were measured by averaging 2 readings while the participant was seated in a chair following a minimum of 5 minutes of rest at the research center. Hypertension was defined as SBP ≥130 mm Hg, DBP ≥80 mm Hg, or self-reported use of antihypertensive medications. Pulse pressure (PP) was calculated by taking the difference between SBP and DBP (SBP-DBP). Mean arterial pressure (MAP) was calculated as the integrated mean of the calibrated brachial pressure waveform at the time of the arterial tonometry test at examination 3. Fasting total cholesterol, high-density lipoprotein cholesterol, triglycerides, and blood glucose were also obtained during examination 3 at the research center.

Statistical Analyses

We reported mean and SD for continuous variables with approximately normal distributions, median and IQR for continuous variables with skewed distributions, and frequency and proportion for categorical variables. We conducted linear regressions to investigate the relations between arterial stiffness traits (predictors) and home BPV as estimated by ARV (outcomes). Prior to regression analysis, CFPWV was inverse-transformed to reduce heteroscedasticity and skewness and was then multiplied by –1000 to convert the units to milliseconds per meter and restore directionality. Therefore, the transformed CFPWV was expressed as –1000/CFPWV. To facilitate comparison and interpretation, all predictors and outcomes were scaled to unit SD in the regression analysis.

For each of the arterial stiffness traits (CFPWV, FWA, and RC), the association with home BPV was evaluated using separate linear models for ARV indices derived from home SBP and home DBP. In the base models, we performed linear regressions adjusting for age, age-squared, and sex. We adjusted for age-squared because previous studies observed that many stiffness measures and BP measures showed nonlinear age relations [22]. In the multivariable models, all base models were further adjusted for the following covariates: BMI, height, heart rate, total cholesterol, high-density lipoprotein cholesterol, triglycerides, lipid-lowering treatment, fasting glucose, diabetes, current smoking, and antihypertensive medication. The multivariable model helps us determine if there is a residual association of stiffness after adjusting for the other partially downstream and partially independent or upstream effects. In addition, in secondary analysis, to account for any pressure dependence of stiffness variables, we investigated how MAP influenced the associations between arterial stiffness traits and ARV of home BP by further adjusting for MAP in the multivariable model.

All statistical analyses were conducted using R (version 4.0; R Foundation for Statistical Computing). We used 2-tailed $P<.05$ for significance.

Ethical Considerations

All study participants provided informed consent. The eFHS and FHS protocols were approved by the institutional review board at the Boston University Medical Center (H-36586 and H-32132). We confirm that we have the permission to use the data.

Results

Participant Characteristics

Characteristics of the study sample are summarized in Table 1. Our study sample consisted of 857 participants who were middle-aged on average, with a moderate prevalence of hypertension. Average SBP and DBP were within the normal range, whereas approximately 1 in 5 participants reported taking antihypertensive medications. Participants were overweight on average, whereas the prevalence of smoking and diabetes was low. Compared with the BP readings at examination 3, the average home SBP during eFHS over 1-year of follow-up was slightly higher, and the average home DBP was similar. Compared with the FHS attendees who did not enroll in eFHS and the eFHS participants ineligible for our final analysis, the eFHS participants in our study sample were generally healthier, more likely to be female, and had a lower prevalence of risk factors. In addition, the final study sample had a lower mean CFPWV, larger mean RC, and lower mean SBP compared with the FHS participants who did not enroll in eFHS. However, compared with the ineligible eFHS participants, while our final study sample also had a larger mean RC, there were no significant differences in terms of CFPWV and SBP (Multimedia Appendix 3).

Table 1. Characteristics of the study sample (N=857).

Characteristics or covariates	Variables
At the time of research examination 3	
Age (years), mean (SD)	54 (9)
Sex (female), n (%)	508 (59)
BMI (kg/m^2), mean (SD)	27.6 (4.79)
Race and ethnicity, n (%)	
Asian	17 (2)
Black	18 (2.1)
Hispanic	19 (2.2)
White	786 (91.7)
Other	17 (2)
Height (inches), mean (SD)	66.6 (3.52)
Total cholesterol (mg/dL), mean (SD)	190 (35.8)
High-density lipoprotein cholesterol (mg/dL), mean (SD)	62.0 (19.9)
Triglycerides (mg/dL), median (IQR)	87 (65-104)
Fasting blood glucose (mg/dL), mean (SD)	97.9 (17.9)
Heart rate (bpm), mean (SD)	58 (9.08)
Antihypertensive use, n (%)	189 (22.1)
Lipid lowering treatment, n (%)	194 (22.6)
Current smoking, n (%)	36 (4.2)
Diabetes mellitus, n (%)	49 (5.7)
Hypertension, n (%)	405 (47)
SBP[a] (mm Hg), mean (SD)	119 (14.1)
DBP[b] (mm Hg), mean (SD)	76 (8.5)
Mean arterial pressure (mm Hg), mean (SD)	92 (10.8)
Arterial tonometry measures	
Carotid-femoral pulse wave velocity (m/second), mean (SD)	7.79 (1.79)
Forward pressure wave amplitude (mm Hg), mean (SD)	47.6 (12.1)
Reflection coefficient, mean (SD)	0.39 (0.07)
Digital home BP[c] during eFHS[d] follow-up	
Follow-up weeks, median (IQR)	47 (21-52)
Number of BP readings, median (IQR)	23 (10-47)
Average SBP (mm Hg), mean (SD)	122 (12.3)
Average DBP (mm Hg), mean (SD)	76 (8.2)
ARV[e] of SBP (mm Hg), mean (SD)	8.61 (3.34)
ARV of DBP (mm Hg), mean (SD)	5.50 (2.22)

[a]SBP: systolic blood pressure.

[b]DBP: diastolic blood pressure.

[c]BP: blood pressure.

[d]eFHS: electronic Framingham Heart Study.

[e]ARV: average real variability.

Association Between Arterial Stiffness Traits and Home BP Variability

We observed that higher CFPWV, FWA, and RC were associated with higher ARV of week-to-week home SBP adjusting for sex, age, and age-squared in the base models (Figure 1). For example, we observed that 1 SD increments in FWA were associated with 0.19 SD increments in the ARV of home SBP. The association of FWA with ARV of home SBP was attenuated but persisted in the multivariable model that adjusted for additional covariates, albeit with a 16% reduction in the magnitude of association. The association between RC and ARV of home SBP was strengthened after including additional covariates in the multivariable model; however, the association between CFPWV and ARV of home SBP was attenuated and became nonsignificant. After further adjusting for MAP, the associations of FWA and RC with ARV of SBP were robust (attenuation: 44% and 43%, respectively; Multimedia Appendix 4).

Next, we performed association analyses between arterial stiffness traits and ARV of week-to-week home DBP (Figure 2). Higher CFPWV and FWA were associated with higher ARV of home DBP in the base models. However, we found no evidence of an association of RC with ARV of home DBP in the base models. The association of FWA with ARV of DBP persisted after adjustment for additional covariates (attenuation: 33%). Higher RC was associated with higher ARV of home DBP with a larger effect estimate, and CFPWV was no longer associated with ARV of home DBP in the multivariable model. When further adjusting for MAP, the directionality of the association between CFPWV and ARV of home DBP was reversed, resulting in higher CFPWV associated with lower ARV of DBP. Neither FWA nor RC was associated with ARV of DBP in the model that further considered MAP (Multimedia Appendix 5).

Figure 1. Association of arterial stiffness traits with ARV of home SBP in the base model and multivariable model. Covariates in base models include sex, age, and age squared. Covariates in the multivariable models include sex, age, age squared, BMI, height, heart rate, total cholesterol, high-density lipoprotein cholesterol, triglycerides, lipid-lowering treatment, fasting glucose, diabetes, current smoking, and antihypertensive medication. ARV: average real variability; CFPWV: carotid-femoral pulse wave velocity; FWA: forward pressure wave amplitude; RC: reflection coefficient; SBP: systolic blood pressure.

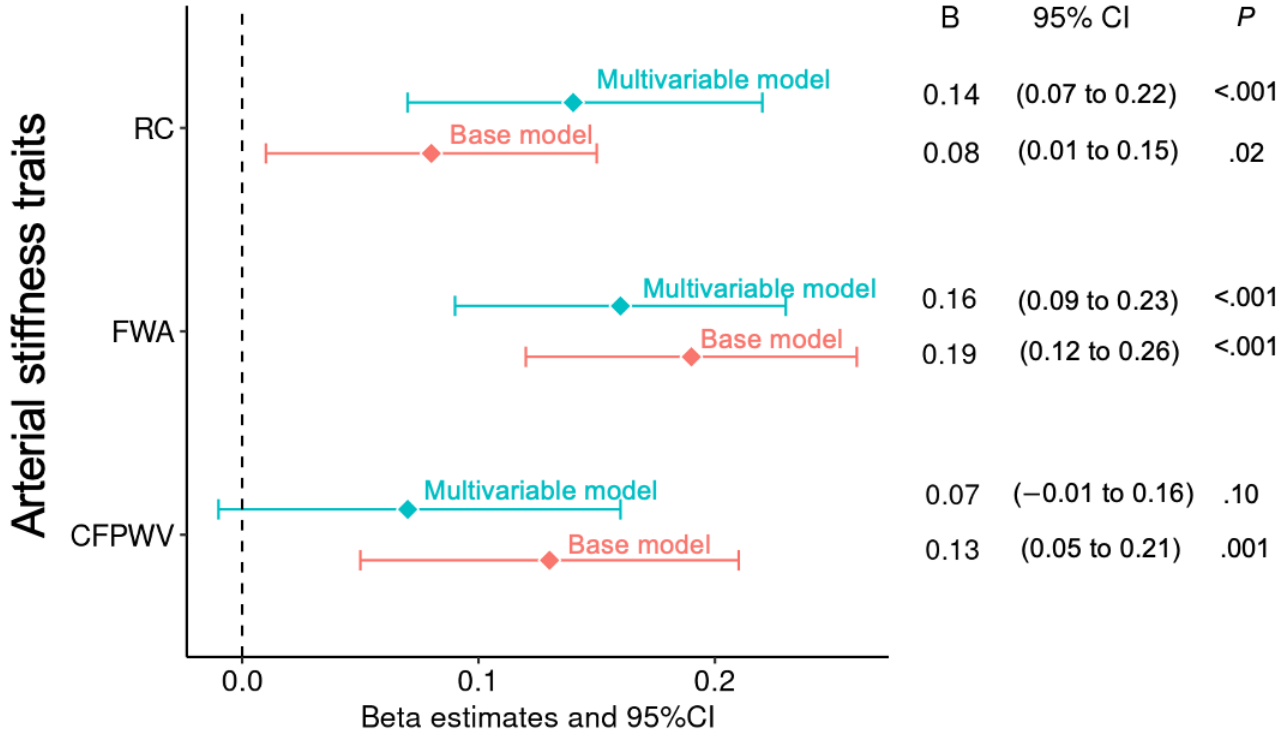
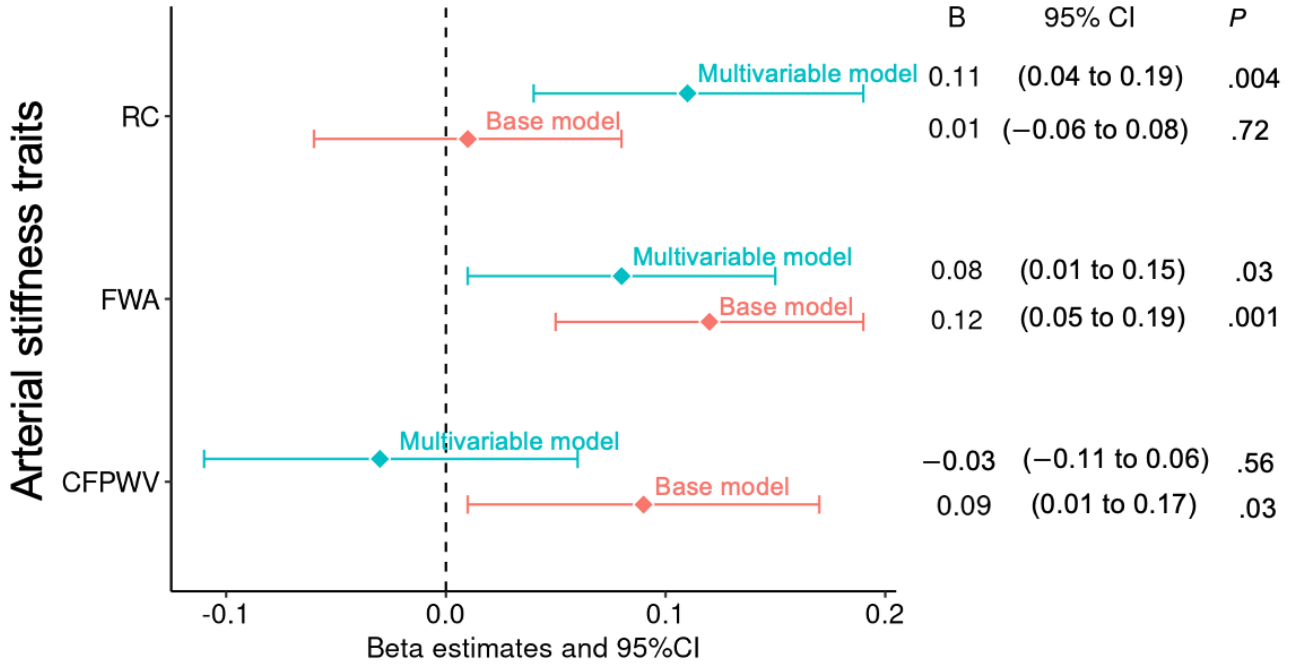


Figure 2. Association of arterial stiffness traits with ARV of home DBP in the base model and multivariable model. Covariates in base models include sex, age, and age squared. Covariates in the multivariable models include sex, age, age squared, BMI, height, heart rate, total cholesterol, high-density lipoprotein cholesterol, triglycerides, lipid-lowering treatment, fasting glucose, diabetes, current smoking, and antihypertensive medication. ARV: average real variability; CFPWV: carotid-femoral pulse wave velocity; DBP: diastolic blood pressure; FWA: forward pressure wave amplitude; RC: reflection coefficient.



Discussion

We investigated the association of arterial stiffness measures with mid- to long-term BPV defined as ARV of week-to-week home SBP and DBP collected using a digital BP cuff for up to 1-year follow-up among middle-aged to older community-dwelling adults with and without hypertension. Higher FWA and RC were associated with higher mid- to long-term ARV of home SBP after adjustment for important covariates (eg, antihypertensive use), while CFPWV was not associated with ARV of home SBP. Similarly, both FWA and RC exhibited positive associations with ARV of DBP in the multivariable models, while there was no evidence of an association of CFPWV with ARV of DBP. After further adjusting for MAP in the secondary analysis, the associations of FWA and RC with ARV of SBP were weakened but persisted compared with multivariable models without accounting for MAP, and we also observed an association between higher CFPWV and lower mid- to long-term ARV of DBP. This negative association relates to the opposing effects of MAP and CFPWV on DBP. An increase in MAP tends to increase DBP, while the MAP-related increase in CFPWV tends to increase PP and therefore decrease DBP [29]. As a result, when the aorta is compliant, changes in MAP will be directly reflected in commensurate changes in DBP. However, as the aorta stiffens, the effects of changes in MAP on DBP will be opposed by changes in PP, resulting in reduced DBP variability. Our findings suggest that measures of higher aortic stiffness and wave reflection were associated with week-to-week variation of BP in a home-based setting.

We used ARV due to its advantages over conventional indices like SD and CV. In comparison, ARV is an average of the absolute differences between consecutive BP measurements. It is more sensitive to the individual BP measurement sequence and may be a better index to represent short-term, reading-to-reading changes. For instance, a steady change (eg, 140, 130, 120, and 110) versus a more chaotic change (eg, 140, 120, 130, and 110) in BP will have the same mean, SD, and CV but a different ARV. Therefore, the ARV considers the order of the measurements and differences in consecutive BP measurements and therefore may be able to better reflect differences between steady change versus more dynamic change.

Central artery stiffness contributes to the pathogenesis of hypertension and the risk for target organ damage in the heart, kidneys, and brain. In contrast, while wave reflection can add to the load on the heart, it may be protective in the periphery by limiting the potentially harmful pulsatile energy transmitted to target organs [14,30]. Among Japanese adults with at least 1 risk factor for cardiovascular disease (CVD), day-to-day home BPV was associated with CVD events in adults with higher baseline arterial stiffness suggesting that arterial stiffness contributes to the association between home BPV and CVD risk [13] observed in a number of studies [6,8,31]. In the population-based Maastricht Study focusing on type 2 diabetes, 7-day systolic BPV was associated with aortic stiffness [32]. Similarly, a study of middle-aged Korean adults with high normal BP or hypertension, identified a significant relationship between home SBP variability and arterial stiffness. In that study, home BP measurements also occurred over 7 consecutive days. Our study extends these findings to home BP measurements taken week-to-week over a longer time horizon

of up to 1 year and observed an association of arterial stiffness measures with mid- to long-term ARV of home BP.

The underlying pathophysiology of systolic versus diastolic BPV has been posited to be different [33]. The main distinction between SBP and DBP variability lies in the differing effects of PP variability and MAP variability. PP variability and MAP variability have additive effects on SBP and offsetting effects on DBP. The ARV of SBP correlates with arterial stiffness as observed in our study while ARV of DBP may be more related to endothelial dysfunction and impaired autonomic function [34].

Out-of-office home BP self-monitoring is a strategy that can be achieved in the community and in low-resource areas to improve hypertension awareness, treatment, and control and is supported by data from the International Databases on Ambulatory and Home Blood Pressure in Relation to Cardiovascular Outcomes [35] and endorsed by the American Heart Association and American College of Cardiology [1]. In addition, the ability to measure BP at home offers the patient the convenience of avoiding some office-based visits, empowers the patient to take multiple measurements over a longer period of time, and can improve engagement with BP management resulting in lower BP [36,37]. Using data from the National Health and Nutrition Examination Survey, less than 50% of people with known hypertension engaged in home BP monitoring at least monthly, leaving more work to be done [38]. With the rising use of mobile phones in the United States across diverse populations (White, Black, Hispanic, urban, and rural) [39], the extension of home BP measurements to digital measurements as in our study may permit the transmission of BP measurements to the health care team, development of educational visualization tools to enhance understanding of the BP measurement, and the use of nudges to encourage an individual to take BP measurements [40].

Our study had several strengths including the well-characterized community-based sample, the arterial stiffness measures obtained with a standardized protocol, and the tracking of BPs in the home setting using a digital device over up to 1 year. Higher SBP variability (visit-to-visit) is associated with adverse CVD outcomes in adults with optimal BP levels irrespective of underlying hypertension [12]. Therefore, it is critical to include individuals across the BP spectrum in studies that investigate BPV and other risk factors for CVD. Our study included individuals across a broad spectrum of BP levels, including those with and without underlying hypertension. This comprehensive inclusion enabled us to contribute valuable insights into understanding the potential biological mechanism that leads to increased risk of CVD with elevated BPV.

Our study also had some limitations that merit comment. First, more than half of the eFHS participants were excluded from our final sample, which led to differences in participant demographics between the study sample and the eFHS participants ineligible for our final analysis. The exclusion process potentially introduced a selection bias and limited the generalization of our findings to individuals of more diverse backgrounds. However, it is important to note that these 2 groups did not show significant differences in the prevalence of hypertension, reported use of antihypertensive medications, or measures of arterial stiffness, except for RC. This similarity may help mitigate the impact of selection bias to some extent. Second, the study is observational and cannot infer a causal relationship. Third, home BP measurements were taken by participants once per week at about the same time of the day each week. This schedule limited our ability to collect more frequent BP data, consequently restricting our ability to investigate BPV within a day or across different days within the same week. Advances in technology include wearables that measure BP at the wrist, which can provide more frequent data and a more convenient approach to collecting BP measurements than our method. However, more data are needed to determine the accuracy and usefulness of these devices. In addition, we did not correct for multiple statistical tests, which may lead to an inflated type I error. However, traditional correction techniques such as the Bonferroni correction would be overly conservative in our case due to correlated traits. Notably, upon applying multiple testing corrections, the results in our primary analysis remained largely unaffected, except for the association between ARV of home DBP and FWA, which subsequently lost its statistical significance.

In conclusion, in our middle-aged to older adult community-based sample, higher FWA and RC were associated with higher mid- to long-term ARV of home SBP and ARV of home DBP. Our findings suggest that higher aortic stiffness, as assessed by FWA, and greater relative wave reflection, as assessed by the global RC, may increase the week-to-week variation of home-based BP. The association of arterial stiffness with ARV of home BP may be a biological mechanism contributing to the increased risk of CVD associated with home BPV. Future work is needed to determine the potential beneficial effects on CVD outcomes of targeted attempts to reduce home BPV. The importance of defining the contribution to cardiovascular risk and outcomes of not only magnitude and duration but also variability of risk factors, including BP, is being recognized and requires additional work.

Acknowledgments

This study was supported by a grant (R01HL141434); the Framingham Heart Study (contract 75N92019D00031 and RWJ 74624); and the National Heart, Lung, and Blood Institute (contracts N01-HC-25195, HHSN268201500001l, and 75N92019D00031; [RSV]; R01-DK-080739 [RSV], R01-HL-107385, 1R01HL126136-01A1, HL93328, HL142983, HL143227, and HL131532 [RSV and GFM]; and 1RO1-HL-70100, R01HL092577, 2U54HL120163, and 1R01AG066010 [EJB]).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request and will be available at public repositories (DbGAP and BioLINNC).

Conflicts of Interest

GFM is the owner of Cardiovascular Engineering, Inc, a company that designs and manufactures devices that measure vascular stiffness. The company uses these devices in clinical trials that evaluate the effects of diseases and interventions on vascular stiffness. GFM also serves as a consultant to and receives grants and honoraria from Novartis, Merck, Bayer, Servier, Philips, and deCODE genetics. DDM has received research support from Apple Inc, Bristol-Myers Squibb, Boehringer-Ingelheim, Pfizer, Flexcon, Samsung, Philips Health care, and Biotronik, and has received consultancy fees from Heart Rhythm Society, Bristol-Myers Squibb, Pfizer, Flexcon, Boston Biomedical Associates, and Rose Consulting. DDM also declares financial support for serving on the Steering Committee for the GUARD-AF study (NCT04126486) and the Advisory Committee for the Fitbit Heart Study (NCT04176926). NMH has received funding from the American Heart Association and the National Institutes of Health.

Multimedia Appendix 1
The timeline of data collection for arterial tonometry measures, as well as weekly home BP readings.
[PNG File , 59 KB - cardio_v8i1e54801_app1.png]

Multimedia Appendix 2
Flowchart of sample selection.
[PNG File , 219 KB - cardio_v8i1e54801_app2.png]

Multimedia Appendix 3
Comparison between three groups of samples.
[DOCX File , 19 KB - cardio_v8i1e54801_app3.docx]

Multimedia Appendix 4
Association of arterial stiffness traits with ARV of home SBP in the base model, multivariable model, and multivariable model that adjusted for MAP. Covariates in base models include sex, age and age squared. Covariates in the multivariable models include sex, age, age squared, BMI, height, heart rate, total cholesterol, high-density lipoprotein cholesterol, triglycerides, lipid-lowering treatment, fasting glucose, diabetes, current smoking, and anti-hypertensive medication. The multivariable model with MAP was further adjusted for MAP in addition to the variables in the multivariable model. ARV: average real variability; MAP: mean arterial pressure; RC: reflection coefficient; CFPWV: carotid femoral pulse wave velocity; FWA: forward pressure wave amplitude; SBP: systolic blood pressure.
[PNG File , 149 KB - cardio_v8i1e54801_app4.png]

Multimedia Appendix 5
Association of arterial stiffness traits with ARV of home DBP in the base model, multivariable model, and multivariable model that adjusted for MAP. Covariates in base models include sex, age and age squared. Covariates in the multivariable models include sex, age, age squared, BMI, height, heart rate, total cholesterol, high-density lipoprotein cholesterol, triglycerides, lipid-lowering treatment, fasting glucose, diabetes, current smoking, and anti-hypertensive medication. The multivariable model with MAP was further adjusted for MAP in addition to the variables in the multivariable model. ARV: average real variability; MAP: mean arterial pressure; RC: reflection coefficient; CFPWV: carotid femoral pulse wave velocity; FWA: forward pressure wave amplitude; DBP: diastolic blood pressure.
[PNG File , 146 KB - cardio_v8i1e54801_app5.png]

References

1. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Himmelfarb CD, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. Hypertension 2018;71(6):e13-e115 [FREE Full text] [doi: 10.1161/HYP.0000000000000065] [Medline: 29133356]
2. Niiranen TJ, Hänninen MR, Johansson J, Reunanen A, Jula AM. Home-measured blood pressure is a stronger predictor of cardiovascular risk than office blood pressure: the Finn-Home study. Hypertension 2010;55(6):1346-1351 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.109.149336] [Medline: 20385970]

3. Parati G, Bilo G, Kollias A, Pengo M, Ochoa JE, Castiglioni P, et al. Blood pressure variability: methodological aspects, clinical relevance and practical indications for management—a European Society of Hypertension position paper *. J Hypertens 2023;41(4):527-544 [FREE Full text] [doi: 10.1097/HJH.0000000000003363] [Medline: 36723481]
4. Smith TO, Sillito JA, Goh CH, Abdel-Fattah AR, Einarsson A, Soiza RL, et al. Association between different methods of assessing blood pressure variability and incident cardiovascular disease, cardiovascular mortality and all-cause mortality: a systematic review. Age Ageing 2020;49(2):184-192 [FREE Full text] [doi: 10.1093/ageing/afz178] [Medline: 31985773]
5. Juhanoja EP, Niiranen TJ, Johansson JK, Puukka PJ, Thijs L, Asayama K, et al. Outcome-driven thresholds for increased home blood pressure variability. Hypertension 2017;69(4):599-607 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.116.08603] [Medline: 28193705]
6. Johansson JK, Niiranen TJ, Puukka PJ, Jula AM. Prognostic value of the variability in home-measured blood pressure and heart rate: the Finn-Home Study. Hypertension 2012;59(2):212-218 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.178657] [Medline: 22215704]
7. Asayama K, Kikuya M, Schutte R, Thijs L, Hosaka M, Satoh M, et al. Home blood pressure variability as cardiovascular risk factor in the population of Ohasama. Hypertension 2013;61(1):61-69 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.00138] [Medline: 23172933]
8. Kikuya M, Ohkubo T, Metoki H, Asayama K, Hara A, Obara T, et al. Day-by-day variability of blood pressure and heart rate at home as a novel predictor of prognosis: the Ohasama study. Hypertension 2008;52(6):1045-1050 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.107.104620] [Medline: 18981332]
9. Matsumoto A, Satoh M, Kikuya M, Ohkubo T, Hirano M, Inoue R, et al. Day-to-day variability in home blood pressure is associated with cognitive decline: the Ohasama study. Hypertension 2014;63(6):1333-1338 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.113.01819] [Medline: 24688128]
10. Stevens SL, Wood S, Koshiaris C, Law K, Glasziou P, Stevens RJ, et al. Blood pressure variability and cardiovascular disease: systematic review and meta-analysis. BMJ 2016;354:i4098 [FREE Full text] [doi: 10.1136/bmj.i4098] [Medline: 27511067]
11. Messerli FH, Hofstetter L, Rimoldi SF, Rexhaj E, Bangalore S. Risk factor variability and cardiovascular outcome: JACC review topic of the week. J Am Coll Cardiol 2019;73(20):2596-2603 [FREE Full text] [doi: 10.1016/j.jacc.2019.02.063] [Medline: 31118154]
12. Liu M, Chen X, Zhang S, Lin J, Wang L, Liao X, et al. Assessment of visit-to-visit blood pressure variability in adults with optimal blood pressure: a new player in the evaluation of residual cardiovascular risk? J Am Heart Assoc 2022;11(9):e022716 [FREE Full text] [doi: 10.1161/JAHA.121.022716] [Medline: 35470678]
13. Ishiyama Y, Hoshide S, Kanegae H, Kario K. Impact of home blood pressure variability on cardiovascular outcome in patients with arterial stiffness: results of the J-HOP study. J Clin Hypertens (Greenwich) 2021;23(8):1529-1537 [FREE Full text] [doi: 10.1111/jch.14327] [Medline: 34288371]
14. Kaess BM, Rong J, Larson MG, Hamburg NM, Vita JA, Levy D, et al. Aortic stiffness, blood pressure progression, and incident hypertension. JAMA 2012;308(9):875-881 [FREE Full text] [doi: 10.1001/2012.jama.10503] [Medline: 22948697]
15. Mitchell GF, Hwang SJ, Vasan RS, Larson MG, Pencina MJ, Hamburg NM, et al. Arterial stiffness and cardiovascular events: the Framingham Heart Study. Circulation 2010;121(4):505-511 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.886655] [Medline: 20083680]
16. Vasan RS, Pan S, Xanthakis V, Beiser A, Larson MG, Seshadri S, et al. Arterial stiffness and long-term risk of health outcomes: the Framingham Heart Study. Hypertension 2022;79(5):1045-1056 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.121.18776] [Medline: 35168368]
17. Tedla YG, Yano Y, Carnethon M, Greenland P. Association between long-term blood pressure variability and 10-year progression in arterial stiffness: the multiethnic study of atherosclerosis. Hypertension 2017;69(1):118-127 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.116.08427] [Medline: 27849565]
18. Schillaci G, Bilo G, Pucci G, Laurent S, Macquin-Mavier I, Boutouyrie P, et al. Relationship between short-term blood pressure variability and large-artery stiffness in human hypertension: findings from 2 large databases. Hypertension 2012;60(2):369-377 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.112.197491] [Medline: 22753222]
19. Park JS, Shin JH, Park JB, Choi DJ, Youn HJ, Park CG, et al. Relationship between arterial stiffness and variability of home blood pressure monitoring. Medicine (Baltimore) 2020;99(30):e21227 [FREE Full text] [doi: 10.1097/MD.0000000000021227] [Medline: 32791697]
20. Goldberg EM, Levy PD. New approaches to evaluating and monitoring blood pressure. Curr Hypertens Rep 2016;18(6):49 [FREE Full text] [doi: 10.1007/s11906-016-0650-9] [Medline: 27137524]
21. Topouchian J, Agnoletti D, Blacher J, Youssef A, Chahine MN, Ibanez I, et al. Validation of four devices: Omron M6 comfort, Omron HEM-7420, withings BP-800, and polygreen KP-7670 for home blood pressure measurement according to the European Society of Hypertension International Protocol. Vasc Health Risk Manag 2014;10:33-44 [FREE Full text] [doi: 10.2147/VHRM.S53968] [Medline: 24476688]
22. Mitchell GF, Parise H, Benjamin EJ, Larson MG, Keyes MJ, Vita JA, et al. Changes in arterial stiffness and wave reflection with advancing age in healthy men and women: the Framingham Heart Study. Hypertension 2004;43(6):1239-1245 [FREE Full text] [doi: 10.1161/01.HYP.0000128420.01881.aa] [Medline: 15123572]

23. Mitchell GF, Wang N, Palmisano JN, Larson MG, Hamburg NM, Vita JA, et al. Hemodynamic correlates of blood pressure across the adult age spectrum: noninvasive evaluation in the Framingham Heart Study. Circulation 2010;122(14):1379-1386 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.914507] [Medline: 20855656]
24. Townsend RR, Wilkinson IB, Schiffrin EL, Avolio AP, Chirinos JA, Cockcroft JR, et al. Recommendations for improving and standardizing vascular research on arterial stiffness: a scientific statement from the American Heart Association. Hypertension 2015;66(3):698-722 [FREE Full text] [doi: 10.1161/HYP.0000000000000033] [Medline: 26160955]
25. Mitchell G, van Buchem MA, Sigurdsson S, Gotal JD, Jonsdottir MK, Kjartansson Ó, et al. Arterial stiffness, pressure and flow pulsatility and brain structure and function: the age, gene/environment susceptibility--Reykjavik study. Brain 2011;134(Pt 11):3398-3407 [FREE Full text] [doi: 10.1093/brain/awr253] [Medline: 22075523]
26. Sardana M, Lin H, Zhang Y, Liu C, Trinquart L, Benjamin EJ, et al. Association of habitual physical activity with home blood pressure in the Electronic Framingham Heart Study (eFHS): cross-sectional study. J Med Internet Res 2021;23(6):e25591 [FREE Full text] [doi: 10.2196/25591] [Medline: 34185019]
27. Mena L, Pintos S, Queipo NV, Aizpúrua JA, Maestre G, Sulbarán T. A reliable index for the prognostic significance of blood pressure variability. J Hypertens 2005;23(3):505-511. [doi: 10.1097/01.hjh.0000160205.81652.5a] [Medline: 15716690]
28. Mena LJ, Maestre GE, Hansen TW, Thijs L, Liu Y, Boggia J, et al. How many measurements are needed to estimate blood pressure variability without loss of prognostic information? Am J Hypertens 2014;27(1):46-55 [FREE Full text] [doi: 10.1093/ajh/hpt142] [Medline: 23955605]
29. Mitchell GF. Imaging tools in cardiovascular research. Clin Transl Sci 2009:105-121. [doi: 10.1016/b978-0-12-373639-0.00007-8]
30. Mitchell GF. Arterial stiffness: insights from Framingham and Iceland. Curr Opin Nephrol Hypertens 2015;24(1):1-7 [FREE Full text] [doi: 10.1097/MNH.0000000000000092] [Medline: 25470012]
31. Hoshide S, Yano Y, Mizuno H, Kanegae H, Kario K. Day-by-day variability of home blood pressure and incident cardiovascular disease in clinical practice: the J-HOP Study (Japan Morning Surge-Home Blood Pressure). Hypertension 2018;71(1):177-184 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.117.10385] [Medline: 29133364]
32. Zhou TL, Henry RMA, Stehouwer CDA, van Sloten TT, Reesink KD, Kroon AA. Blood pressure variability, arterial stiffness, and arterial remodeling. Hypertension 2018;72(4):1002-1010 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.118.11325] [Medline: 30354707]
33. Sheikh AB, Sobotka PA, Garg I, Dunn JP, Minhas AMK, Shandhi MMH, et al. Blood pressure variability in clinical practice: past, present and the future. J Am Heart Assoc 2023;12(9):e029297 [FREE Full text] [doi: 10.1161/JAHA.122.029297] [Medline: 37119077]
34. Parati G, Stergiou GS, Dolan E, Bilo G. Blood pressure variability: clinical relevance and application. J Clin Hypertens (Greenwich) 2018;20(7):1133-1137 [FREE Full text] [doi: 10.1111/jch.13304] [Medline: 30003704]
35. Asayama K, Stolarz-Skrzypek K, Yang WY, Hansen TW, Brguljan-Hitij J, Odili AN, et al. What did we learn from the international databases on ambulatory and home blood pressure in relation to cardiovascular outcome? Hypertens Res 2023;46(4):934-949 [FREE Full text] [doi: 10.1038/s41440-023-01191-4] [Medline: 36737461]
36. McManus RJ, Mant J, Haque MS, Bray EP, Bryan S, Greenfield SM, et al. Effect of self-monitoring and medication self-titration on systolic blood pressure in hypertensive patients at high risk of cardiovascular disease: the TASMIN-SR randomized clinical trial. JAMA 2014;312(8):799-808 [FREE Full text] [doi: 10.1001/jama.2014.10057] [Medline: 25157723]
37. Herrigel DJ, Willett LR. In hypertension with high CV risk, BP self-monitoring and medication self-titration reduced BP at 12 months. Ann Intern Med 2015;162(2):JC7. [doi: 10.7326/ACPJC-2015-162-2-007] [Medline: 25599367]
38. Ostchega Y, Berman L, Hughes JP, Chen TC, Chiappa MM. Home blood pressure monitoring and hypertension status among US adults: the National Health and Nutrition Examination Survey (NHANES), 2009-2010. Am J Hypertens 2013;26(9):1086-1092 [FREE Full text] [doi: 10.1093/ajh/hpt054] [Medline: 23604493]
39. Mobile fact sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/#mobile-phone-ownership-over-time [accessed 2022-01-26]
40. Ding EY, Pathiravasan CH, Schramm E, Borrelli B, Liu C, Trinquart L, et al. Design, deployment, and usability of a mobile system for cardiovascular health monitoring within the electronic Framingham Heart Study. Cardiovasc Digit Health J 2021;2(3):171-178 [FREE Full text] [doi: 10.1016/j.cvdhj.2021.04.001] [Medline: 35265906]

Abbreviations

ARV: average real variability
BP: blood pressure
BPV: blood pressure variability
CFPWV: carotid-femoral pulse wave velocity
CV: coefficient of variation
CVD: cardiovascular disease
DBP: diastolic blood pressure
eFHS: electronic Framingham Heart Study

FHS: Framingham Heart Study
FWA: forward pressure wave amplitude
Gen 3: third generation
MAP: mean arterial pressure
PP: pulse pressure
RC: reflection coefficient
SBP: systolic blood pressure

Edited by A Mavragani; submitted 22.11.23; peer-reviewed by N Kakaletsis; comments to author 20.12.23; revised version received 26.01.24; accepted 29.02.24; published 08.04.24.

Please cite as:
Wang X, Zhang Y, Pathiravasan CH, Ukonu NC, Rong J, Benjamin EJ, McManus DD, Larson MG, Vasan RS, Hamburg NM, Murabito JM, Liu C, Mitchell GF
Association of Arterial Stiffness With Mid- to Long-Term Home Blood Pressure Variability in the Electronic Framingham Heart Study: Cohort Study
JMIR Cardio 2024;8:e54801
URL: https://cardio.jmir.org/2024/1/e54801
doi:10.2196/54801
PMID:38587880

Original Paper

Accurate Modeling of Ejection Fraction and Stroke Volume With Mobile Phone Auscultation: Prospective Case-Control Study

Martin Huecker[1], MD; Craig Schutzman[1], MD; Joshua French[1], MD; Karim El-Kersh[2], MD; Shahab Ghafghazi[1], MD; Ravi Desai[3], MD; Daniel Frick[1], MD; Jarred Jeremy Thomas[1], MD

[1]Department of Emergency Medicine, University of Louisville, Louisville, KY, United States

[2]Department of Pulmonary and Critical Care Medicine, The University of Arizona, Phoenix, AZ, United States

[3]Lehigh Valley Health Network Cardiology and Critical Care, Allentown, PA, United States

Corresponding Author:
Martin Huecker, MD
Department of Emergency Medicine
University of Louisville
530 South Jackson St.
Louisville, KY, 40202
United States
Phone: 1 5028525689
Email: martin.huecker@louisville.edu

Abstract

Background: Heart failure (HF) contributes greatly to morbidity, mortality, and health care costs worldwide. Hospital readmission rates are tracked closely and determine federal reimbursement dollars. No current modality or technology allows for accurate measurement of relevant HF parameters in ambulatory, rural, or underserved settings. This limits the use of telehealth to diagnose or monitor HF in ambulatory patients.

Objective: This study describes a novel HF diagnostic technology using audio recordings from a standard mobile phone.

Methods: This prospective study of acoustic microphone recordings enrolled convenience samples of patients from 2 different clinical sites in 2 separate areas of the United States. Recordings were obtained at the aortic (second intercostal) site with the patient sitting upright. The team used recordings to create predictive algorithms using physics-based (not neural networks) models. The analysis matched mobile phone acoustic data to ejection fraction (EF) and stroke volume (SV) as evaluated by echocardiograms. Using the physics-based approach to determine features eliminates the need for neural networks and overfitting strategies entirely, potentially offering advantages in data efficiency, model stability, regulatory visibility, and physical insightfulness.

Results: Recordings were obtained from 113 participants. No recordings were excluded due to background noise or for any other reason. Participants had diverse racial backgrounds and body surface areas. Reliable echocardiogram data were available for EF from 113 patients and for SV from 65 patients. The mean age of the EF cohort was 66.3 (SD 13.3) years, with female patients comprising 38.3% (43/113) of the group. Using an EF cutoff of ≤40% versus >40%, the model (using 4 features) had an area under the receiver operating curve (AUROC) of 0.955, sensitivity of 0.952, specificity of 0.958, and accuracy of 0.956. The mean age of the SV cohort was 65.5 (SD 12.7) years, with female patients comprising 34% (38/65) of the group. Using a clinically relevant SV cutoff of <50 mL versus >50 mL, the model (using 3 features) had an AUROC of 0.922, sensitivity of 1.000, specificity of 0.844, and accuracy of 0.923. Acoustics frequencies associated with SV were observed to be higher than those associated with EF and, therefore, were less likely to pass through the tissue without distortion.

Conclusions: This work describes the use of mobile phone auscultation recordings obtained with unaltered cellular microphones. The analysis reproduced the estimates of EF and SV with impressive accuracy. This technology will be further developed into a mobile app that could bring screening and monitoring of HF to several clinical settings, such as home or telehealth, rural, remote, and underserved areas across the globe. This would bring high-quality diagnostic methods to patients with HF using equipment they already own and in situations where no other diagnostic and monitoring options exist.

(JMIR Cardio 2024;8:e57111) doi:10.2196/57111

KEYWORDS

ejection fraction; stroke volume; auscultation; digital health; telehealth; acoustic recording; acoustic recordings; acoustic; mHealth; mobile health; mobile phone; mobile phones; heart failure; heart; cardiac; cardiology; health care costs; audio; echocardiographic; echocardiogram; ultrasonography; echocardiography; accuracy; monitoring; telemonitoring; recording; recordings; ejection; machine learning; algorithm; algorithms

Introduction

Cardiovascular disorders contribute immensely to morbidity and mortality in the United States and worldwide. Heart failure (HF) is defined as "a clinical syndrome with symptoms and/or signs caused by a structural and/or functional cardiac abnormality and corroborated by elevated natriuretic peptide levels and/or objective evidence of pulmonary or systemic congestion" [1]. At least 64.3 million people around the world have HF, with that number expected to increase due to improved health care [2]. HF accounts for 1% to 2% of all hospitalizations in high-income countries and is the top cause of admission for patients older than 65 years of age [2]. The United States spent more than US $30 billion on HF in 2012, with a projected increase to US $69.8 billion by 2030 [2]. The mortality of HF ranges from as low as 2% to 4% per year in those with chronic HF and up to 36.5% in those with acute HF [2]. In the United States, the mandatory federal pay-for-performance Hospital Readmissions Reduction Program targets patients with HF and ties reimbursement to 30-day, all-cause, Medicare, fee-for-service readmissions after initial hospitalization for HF; rates reach as high as 23% in some studies [3].

HF is divided into three categories based on the left ventricular ejection fraction (LVEF): (1) HF with reduced ejection fraction (EF), (2) mildly reduced EF, and (3) preserved EF, with EF ranges of ≤40%, 41% to 49%, and ≥50%, respectively [1,2]. LVEF, the percentage of blood in the left ventricle that exits into the aorta during a cardiac cycle, is determined using various imaging techniques, such as echocardiography, cardiac magnetic resonance imaging, nuclear cardiology, or cardiac catheterization [1,4,5]. Thus, the classification of HF depends on the accurate determination of LVEF using expensive diagnostic methods obtained in outpatient or inpatient settings [6,7]. A study from the United Kingdom found that most new HF cases were diagnosed in inpatient settings despite the presence of symptoms that should have triggered an earlier outpatient evaluation [8]. This is at least partly due to barriers such as the availability of transportation, cost concerns, and access to medical facilities. Millions of potential patients with HF worldwide lack access to even basic medical care and are, therefore, unable to undergo risk assessment for heart disease.

The management of patients diagnosed with HF involves serial testing to detect changes in heart function. The techniques used to measure LVEF and other cardiac parameters (cardiac output, indexed stroke volume, etc) can have significant variability, limiting prognostication and treatment efficacy [5]. Diagnostic tests to determine EF also experience great variability, limiting prognostication and treatment efficacy [5]. Due to the somewhat limiting paradigm of EF categories, more regular use of vital measures such as stroke volume (SV) could delineate patients with HF with more granularity, even having implications for treatment [9]. Telehealth represents a potential mechanism to reduce the rates of 30-day readmission in patients with HF [10]. Patients without access to large hospital systems and diagnostic testing would benefit immensely from a low-cost yet accurate method of determining these parameters. The technology harnessing more than 8 billion global mobile phones could vastly improve health care disparities [11].

This pilot study describes a novel diagnostic technology using audio recordings from a standard mobile phone. Prior publications have sought both invasive and noninvasive means of describing cardiac function, but very few have moved out of research phases to clinical or practical use [12-16]. This study aims to establish a set of markers using complex but reproducible mathematics from mobile phone auscultation data that would enable the determination of EF and SV for HF detection, classification, and monitoring. The goal of this study was to demonstrate the feasibility of creating mobile phone models for the classification of LVEF and SV by matching echocardiographic results to phone recordings.

Methods

Settings and Participants

This is a pilot prospective study of convenience samples of patients presenting to 2 hospital systems for cardiac workups. At site 1, an urban academic center in the Southern United States, study personnel obtained recordings from patients who received inpatient clinical evaluation for cardiac disease. All participants had a transthoracic echocardiogram within 30 days. At site 2, a large community clinical site in the Northeastern United States, patients already scheduled for outpatient transthoracic echocardiogram were enrolled at the time of the study, and recordings were obtained at the same time as the echo. To minimize audible confounding, the team excluded patients with mechanical heart valves. Patients were also excluded if they had a positive SARS-CoV-2 test, were younger than 18 years of age, or were pregnant.

Ethical Considerations

This study was approved by the human participants' research institutional review boards of the University of Louisville (number 20.0605) at both clinical sites. Written informed consent was obtained from all participants, with the specification that data obtained would be used for research. Patients had the freedom to withdraw at any time, including after data collection and analysis. Privacy and confidentiality were protected by storing all data in secure, encrypted locations. At all times, only IRB-approved personnel had access to the stored data. The participants did not receive any compensation for participation. Patients were consented after the completion of any urgent or emergent diagnostic testing or treatment and after evaluation by inpatient physician teams, to ensure that the study would not

delay necessary evaluation or treatment. The study team did not recommend, order, or perform any testing.

Data Collection

The research team obtained demographics and clinical information from the electronic medical record at each site. Echocardiography was obtained by a single laboratory at each clinical site. The coders used the EF from the final interpretation of the echocardiogram report. Data from site 1 were uploaded into CardBox (Box Inc), a web-based, encrypted research cloud space. Data from site 2 were uploaded to password-protected Google Drive. Data included demographics and formal echo results, as well as other data such as cardiac catheterization reports, vascular imaging, and primary admission diagnoses. Clinical data were matched to respective (deidentified) sound recording files using unique identification codes. SV estimates were based on the Teichholz method—not because it was preferred, but because it was available on most echocardiogram reports.

Technology and Analytic Method

In addition to open-source Python (Python Software Foundation)-based software, 2 proprietary software were used in the study. The first is Another Sound Recorder (ASR), a recording app developed by NLL APPS. It allowed all recordings to be made in a standardized format across the various phone brands used in the study. The second is Time Series Dynamics (TSD) software developed by Fleming Scientific. It maps time series observation of systems, such as auscultation and waveform data, into a set of descriptive "features." The mapping relies entirely on models from "dynamics" which, in physics, is the study of motion resulting from force. These "physics-based" features can then be reduced and used as dependent variables in rigorous statistical modeling. The TSD approach, which intends to preserve physical and mathematical rigor throughout the modeling process, eliminates the need for neural networks and makes it possible to work effectively with smaller data sets [17]. The research team has extensive experience with the TSD approach and is currently using it in analogous respiratory mobile phone auscultation studies funded by the National Institutes of Health (NIH) and the Biomedical Advanced Research and Development Authority (BARDA).

Cardiac auscultation acoustics represent primarily the sounds of hemodynamics, which is the movement of blood resulting from forces applied by the heart and vascular system. In traditional auscultation, providers use these acoustics to make inferences about organ and system functionality. The approach in this paper is analogous, except that the acoustics are mapped by dynamics-based models. The approach also differs from more common machine learning approaches to auscultation data processing that typically rely on some combination of frequency domain, linear stochastics, and neural networks.

The research team hypothesizes that, in the classification of hemodynamics, the use of dynamics-based mapping is domain relevant. It is also consistent with published chaos-based and enthalpic-based views of cardiac function [18,19]. In the approach, thousands of dynamics-based features were extracted from the acoustic recordings by TSD software. Selected features were then matched to echocardiogram findings by simple logistic regression [20]. To maintain statistical rigor by avoiding overfitting, the number of features used in the regression was limited to 4, which represents the minimum number of positive or negative testing cases divided by 10. The 4 features were selected by the maximum entropy method that produced 35 similar combinations that were evaluated separately for best performance. Dimensionality reduction was sufficient to eliminate the need for a validation step.

Developed by Fleming Scientific, this proprietary unpublished technology extracts features found in sound recordings from microphones of unmodified mobile phones. By using models from actual physical acoustics, we created algorithms to match echocardiogram findings. The physics models are designed to describe hemodynamics from the acoustic data, thereby making it possible to classify organ functionality directly. The method produces thousands of candidate features for modeling but uses only a few to avoid overfitting. The features were matched to echocardiogram findings by using logistic regression [20]. The approach eliminates the need for neural networks entirely and offers a more rigorous approach to developing artificial intelligence (AI) software.

The research team obtained audio recordings with an assortment of unmodified, nonencased Android mobile phones including LG and Motorola Trac phones and 2 Samsung Galaxy models. The voice recorder was standardized by using ASR, a free open-source app easily installed on any Android product. Both sites used the following ASR settings: WAV format, similar frame speed, mono recording, and no filters or other settings activated. Recordings took place in settings with moderate background noise, such as emergency department rooms, inpatient rooms, and echocardiography labs.

Study personnel obtained the recordings by pressing the microphone lightly into the patient's skin to minimize surface noise. The participants underwent a 20-second recording at the aortic valve area (second intercostal space just to the right of the sternum). The participants were not required to hold their breath. Patients could be in any position for recording, but most were sitting or semirecumbent. The phones were capable of capturing frequencies as low as 10 Hz, which are well below the range of human auscultation perception. The phones were kept in a secure location at each site, for use only by study personnel.

Recordings from patients underwent physics-based analysis to create the features for use in modeling. The features would serve as independent variables while the dependent variables were parameters such as LVEF, determined by diagnostic testing during hospitalization. By matching selected features to the gold-standard parameters from established diagnostic procedures, algorithms were created that enable common phones to reproduce the gold-standard parameters.

The goal of the analysis was to demonstrate the feasibility of creating mobile phone algorithms for the classification of LVEF and SV by matching echocardiographic results to the phone recordings.

TSD differs from machine learning–based AI in that its overarching goal is to deduce the best physics-based models for making algorithms, thereby maintaining rigor as much as possible. The hemodynamics deduced in this study are consistent with published chaos-based and enthalpic-based views of cardiac function [18,19]. However, this study was not designed to provide physiological verification of the deduced physics.

TSD's physics-based approach eliminates the need for neural networks and overfitting strategies entirely, potentially offering advantages in data efficiency, model stability, regulatory visibility, and physical insightfulness [17]. TSD's use of passive signals rather than active signals differs from echocardiogram and most other gold-standard imaging technologies; it uses an analytical foundation designed to describe dynamics directly.

Although the algorithms are based on physics, evaluating them relies on statistical methods consistent with logistic regression analysis. The algorithms were evaluated for the area under the receiver operating curve (AUROC) using the trapezoidal method. Values >0.9 can be interpreted as "excellent," whereas values in the range of 0.8-0.9 can be interpreted as "good" [21]. Sensitivity, specificity, and accuracy were also calculated and presented with confusion matrix values per common practice. The validity of features was also verified by *Z* test>2 criteria in addition to the heuristic argument.

Results

Study Population

In total, 113 patients were enrolled across 2 sites. No recording had to be excluded from the analysis. However, some echocardiogram reports were excluded because of incomplete or inconsistent reporting of EF (n=2) or SV (n=50). From the recent echocardiogram reports, it was possible to match EF findings in 113 patients and estimated SV in 65 patients. For the 113 patients in the EF cohort, the mean age was 66.3 (SD 13.3) years. The cohort consisted of 61.7% (n=70) male patients and 38.3% (n=43) female patients. Regarding race and ethnicity, 77% (n=87) were White, 20.4% (n=23) were Black, and 2.6% (n=3) were Hispanic or Latino. For the 65 patients in the SV cohort, the mean age was 65.5 (SD 12.7) years. The cohort consisted of 66% (n=43) male patients and 34% (n=22) female patients. Regarding race and ethnicity, 74% (n=48) were White and 26% (n=17) were Black. The EF cohort had a mean BMI of 28.3 (SD 6.323) and a mean body surface area (BSA) of 2.03 (SD 0.273). The SV cohort had a mean BMI of 29.3 (SD 6.561) and a mean BSA of 2.05 (SD 0.272).

LVEF Results

The 113-participant EF cohort consisted of 81 participants from site 1 and 32 from site 2. Of note, 57 participants had EF <55% and 56 had an EF>55%. For analysis, the cases were separated into a binary "positive" versus "negative" classification based on the HF disease EF cutoff of 40%. A total of 42 participants with EF ≤40% were designated "positive" in binary classification and they represented 37.2% (n=42) of the cohort; the other 71 (62.8%) participants had EF >40%. The number of features was limited to 4 to avoid overfitting the algorithm. The AUROC was 0.955 ("excellent"), as shown in Table 1. Case separation was also excellent as shown in Figure 1. The EF algorithm accuracy performed similarly across demographics, BSA, and clinical sites (Table 2).

Table 1. Ejection fraction algorithm performance and features.

Cases (N=113)	Model evaluation	Features	*Z* test
True negative (n=68)	AUROC[a] 0.955	1	2.3
False negative (n=2)	Sensitivity 0.952	2	7.2
True positive (n=40)	Specificity 0.958	3	3.9
False positive (n=3)	Accuracy 0.956	4	9.4

[a]AUROC: area under the receiver operating curve.

Figure 1. Actual versus predicted ejection fraction (EF). FN: false negative; FP: false positive; TN: true negative; TP: true positive.

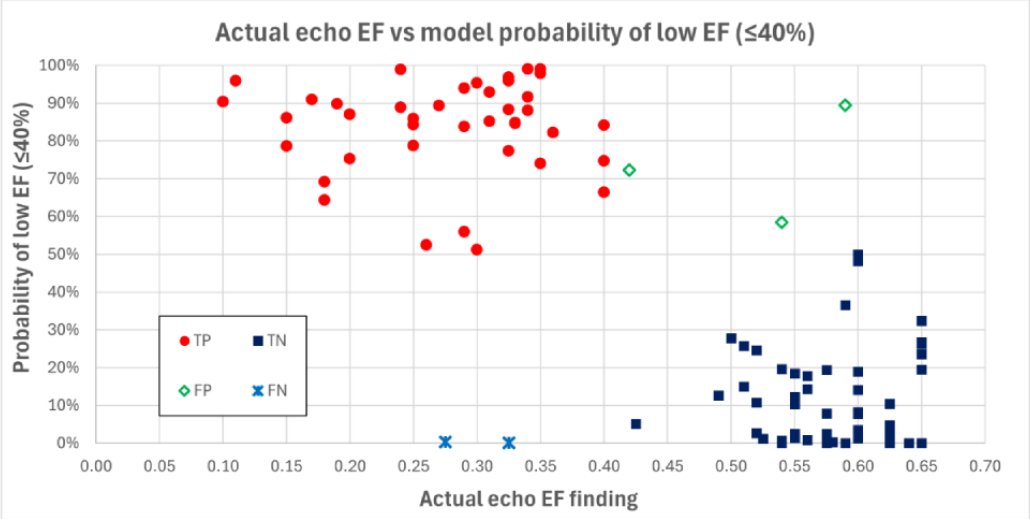


Table 2. EF[a] model had high accuracy across sex, race, BSA[b], and age.

Profile	Accuracy
Sex	
Male	0.94
Female	0.98
Race	
White	0.97
Black or African American	0.95
BSA	
BSA<2.04[c]	0.97
BSA>2.04[c]	0.95
Age (years)	
Younger than 66.3[c]	0.98
Older than 66.3[c]	0.93
Site	
1	0.98
2	0.97

[a]EF: ejection fraction.

[b]BSA: body surface area.

[c]EF sample mean.

SV Results

In all, 65 participants with SV data were all enrolled at site 1. Using a clinically relevant cutoff of <50 mL, 33 (51%) were categorized as positive and 32 (49%) were categorized as negative. For analysis, the number of features was limited to 3 to avoid overfitting the algorithm. Results showed a sensitivity of 100% for the model, with an AUROC of 0.922 (Table 3). Figure 2 illustrates case separation. The SV algorithm accuracy performed similarly across demographics but had a slight drop off in accuracy among patients with higher BSA (Table 4). Acoustics frequencies associated with SV were observed to be higher than those associated with EF and, therefore, were less likely to pass through tissue without distortion.

Table 3. SV[a] algorithm performance and features.

Cases (N=65)	Model evaluation	Features	*Z* test
True negative (n=27)	AUROC[b] 0.922	1	4.0
False negative (n=0)	Sensitivity 1.000	2	2.5
True positive (n=33)	Specificity 0.844	3	3.1
False positive (n=5)	Accuracy 0.923	N/A[c]	N/A

[a]SV: stroke volume.

[b]AUROC: area under the receiver operating curve.

[c]N/A: not applicable.

Figure 2. Actual versus predicted SV. FP: false positive; SV: stroke volume; Teich: Teichholz; TN: true negative; TP: true positive.

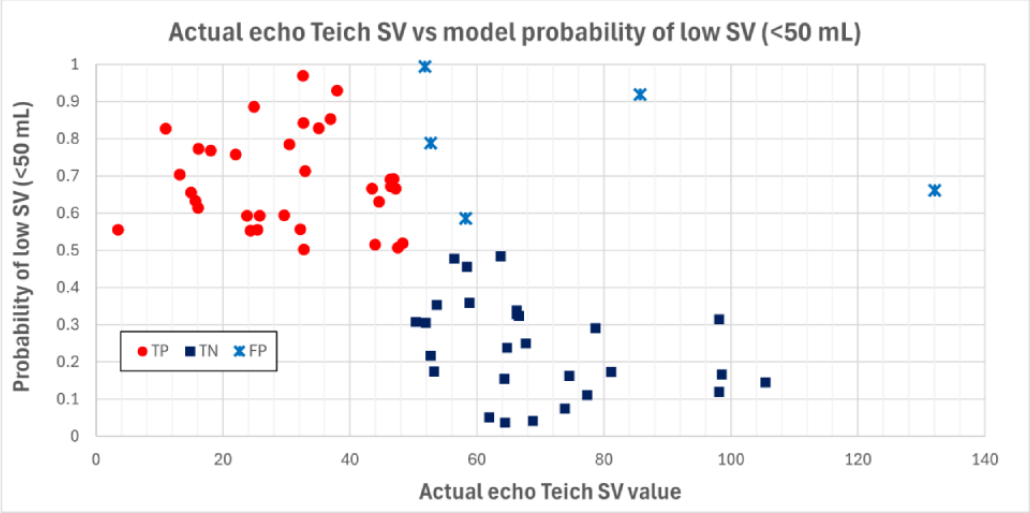


Table 4. SV[a] model had high accuracy across sex, race, BSA[b], and age.

Profile	Accuracy
Sex	
Male	0.93
Female	0.91
Race	
White	0.92
Black or African American	0.94
BSA	
BSA<2.05[c]	0.97
BSA>2.05[c]	0.88
Age (years)	
Younger than 65.5[c]	0.94
Older than 65.5[c]	0.91
Site	
1	0.92

[a]SV: stroke volume.

[b]BSA: body surface area.

[c]SV sample mean.

Discussion

Principal Findings

In this cohort from 2 clinical sites, mobile phone auscultation and dynamics-based modeling allowed accurate detection of low LVEF and SV. These results were obtained using ordinary mobile phones to record from 1 anatomic site with no additional hardware or materials. Prior research suggests that both mobile phones and acoustic recording can assist in HF diagnosis or monitoring; however, no current technologies use basic cellular microphone capability to obtain the acoustic data that can estimate EF or SV. This novel, proprietary unpublished technology has far-reaching potential for screening and management of patients with HF, including the undiagnosed. Perhaps the most obvious use for the technology is telehealth and application to remote and underserved global settings, where even a physical exam by the clinician may not be possible.

Prior work has described technologies that can aid in the monitoring of patients with HF [22-26]. Most technologies use telehealth communications and patient data entry, such as weight, blood pressure, and pulse rate, to risk stratify and monitor disease progress [22,24]. A 2011 Cochrane review established the mortality benefit of telemonitoring in patients with HF [27]. A review by Conway et al [22] identified 4 categories: (1) structured telephone calls; (2) videophone; (3) voice response, which involved the manual input of data using a telephone keypad in response to questions from a computerized voice response system; and (4) telemonitoring. Structured phone calls and telemonitoring showed efficacy in reducing all-cause mortality [22]. Technologies that use true physiologic monitoring require invasive intrathoracic device implantation [28,29], specialized electrocardiography [30], stethoscopes, patches [31,32], or other expensive equipment. Protocols that integrate mobile phones typically use Bluetooth to pair proprietary equipment to a phone in order to transmit data to the care providers [22,25].

Very few described innovations address population screening for HF. A review by Brons et al [33] summarized 99 studies, finding that 100% of algorithms used body weight, 85% used blood pressure, and 61% used heart rate. Bachtiger et al [8] compared 105 patients with low EF to 945 with EF >40% using AI-electrocardiogram (ECG) retrained to interpret a single-lead ECG input. Using a weighted logistic regression from pulmonary and handheld positions, they found an AUROC of 0.91 (95% CI 0.88-0.95), sensitivity of 91.9%, and specificity of 80.2% [8]. One study proposed a method to detect low EF using machine learning or artificial intelligence [34]. Attia et al [35] report on a method using AI-augmented ECG (EKO) to determine the presence of low EF in more than 50,000 patients. The protocol found AUROC, sensitivity, specificity, and accuracy of 0.93%, 86.3%, 85.7%, and 85.7%, respectively. They also found some degree of prediction of future dysfunction: those with a positive AI screen were 4 times more likely to develop ventricular dysfunction in the near future [35].

Shandhi et al [31] compared seismocardiographic data obtained with a wearable sensing patch to objective measurements of pulmonary artery mean pressure and pulmonary capillary wedge pressure following vasodilator infusion during a right heart catheterization, finding reasonable R^2 accuracy (using the Cardiosense technology). These devices use seismocardiological signals in conjunction with ECG signals, thus requiring a hardware device approved by the US Food and Drug Administration (FDA) that must be purchased and maintained. By relying on physics instead of traditional machine learning, a tele-stethoscope does not require the ECG component, making it possible to perform similarly to these more expensive technologies with only an ordinary mobile phone.

One group used computerized acoustic cardiography to detect the third and fourth heart sounds along with systolic time intervals to develop a left ventricular dysfunction index to predict ventricular dysfunction [12]. Their equipment also consisted of an accessory device for a normal ECG machine. Kang et al [36] studied 46 participants to determine the feasibility of phone recordings for detecting heart sounds. Constrained by the presence of 35% of recordings being uninterpretable, the authors found acceptable sensitivity (81%-94%), specificity (79%-100%), positive predictive value (83%-100%), and negative predictive value (82%-92%), with variance depending on which phone was used [36].

Another group tested EF estimation with a novel acoustic-based device (vibration response imaging) that detects low-frequency acoustic signals (10 Hz-70 Hz). The device found sensitivity and specificity around 80%, but the protocol examined requires 36 microphones and a simultaneous ECG [14]. A study using acoustic cardiography in cohorts with and without atrial fibrillation found systolic dysfunction with moderate sensitivity and high specificity (Audicor; Inovise Medical, Inc) [37]. Researchers added sensors to a standard ECG machine to determine 2 systolic parameters: electromechanical activation time and systolic dysfunction index. Another study of the same Audicor device found sensitivity around 80% and specificity in the high 50% range depending on the parameter used [16].

None of these novel approaches show promise for monitoring or diagnosing HF using only mobile phone hardware. Most of the technologies implement proprietary devices and integrate with phones only to transmit data to providers. Tele-stethoscope allows real-time detection of data and rapid transmission of findings directly to clinicians to assist in decision-making. We estimated SV due to its use in approximating cardiac output (SV × heart rate). Noninvasive detection of cardiac output could enhance care for ambulatory and admitted patients. Additionally, SV may represent a parameter that could help distinguish different categories of HF [38].

While this study used research volunteers to obtain the sound recordings, the facile approach allows patients and family members to obtain recordings that can be transmitted with ease using Wi-Fi or cellular signals. This would bring HF diagnosis and monitoring to remote and underserved areas all over the world, to more than 8 billion mobile phones worldwide [11]. Future work will involve matching to other HF diagnostic parameters, such as measures of preserved ejection fraction (early to late diastolic transmitral flow velocity [E/A] to assess diastolic function, and E to early diastolic mitral annular tissue velocity [E/e'] to estimate left ventricular filling pressures) and

pulmonary disease markers (spirometry, chronic obstructive pulmonary disease severity scores, and emphysematous changes on computed tomography imaging). In 1 earlier large-scale human study, this technology was used to match phone acoustics to COVID-19 polymerase chain reaction test results to produce a reliable device for disease detection [39].

Limitations

This work has important limitations. Although relatively small, the sample size was sufficient to demonstrate the feasibility of reproducing echocardiogram EF and SV findings. Additionally, the sample included patients in 2 different cities at 2 different medical centers, 1 inpatient and 1 outpatient. Further studies could center on larger sample sizes and more representative (race, sex, and living areas) recruiting. In a true patient diagnostic model, the best available gold-standard test results, confirmed by diagnosis, would be used rather than echocardiogram reports alone. The enrollment was based on a convenience sample, creating potential selection bias. In the phase 2 study, larger sample sizes will make possible the test or train analysis to demonstrate reproducibility. Larger sample sizes would also make it possible to add more features, if necessary, and reduce the population margin of error.

According to the FDA, a mobile medical app is "a mobile app that incorporates device software functionality that meets the definition of device in section 201(h) of the FD&C Act 11; and either is intended to be used as an accessory to a regulated medical device; or to transform a mobile platform into a regulated medical device" [40]. According to this language, mobile phones and stethoscopes can be considered equivalent. Regarding applicability, the research team views this as a strength rather than a limitation, opening the technology to resource-poor settings all around the world. This would allow fully impromptu data collection in situations where advanced diagnostic equipment is not available and even a physical exam is not possible (telehealth). Phones must be placed directly on the skin and have no motion across the skin, a consideration of importance in future studies where patients will take their measurements. Of note, the fidelity of recordings from this study was not disrupted by background noise; future use in other settings such as ambulance or combat will likely not be limited by ambient noise. Additionally, multiple phone brands were used in the study without any discernible impact on the algorithms.

The comparison of recordings to echocardiogram opens the potential for inaccuracy as transthoracic echocardiogram can have somewhat large margins of error, especially related to EF. The 40% threshold for EF is intended to reduce the rate of false positives. Future work in larger cohorts will allow for a more granular separation of participants. Ongoing work includes recruitment in right heart catheterization and cardiac magnetic resonance imaging patients. Additionally, not all participants at site 1 had the index echocardiogram on the same admission during which acoustic recordings were obtained, but all had the echocardiogram within a 30-day window. Results found no difference in accuracy based on the clinical site or the time of echocardiogram. We did not collect data on the volume status of the participants in the study; acoustic data could potentially vary based on volume status.

It should be noted that no viable features were produced through spectral analysis. One possible explanation is that spectral analysis was unable to manage the nonlinearity of the acoustic signals. Another possible explanation is that it inadvertently created false neighbors among different physical phenomena that happen to share common spectral bands such as low-frequency blood and muscle sounds. Purely from a physics point of view, the features can be interpreted as representing descriptions of fluid and thermodynamics. Although the features used in the modeling are "dynamics-based," and apparently useful in the modeling, their exact physiological interpretation is unknown. At this stage, all that can be said about these features is that they represent some novel interpretation of hemodynamics as "dynamics."

Conclusions

Cardiovascular disease and in particular HF continues to have high morbidity, mortality, and cost worldwide. In this pilot cohort of patients from 2 clinical sites in 2 different cities, passive acoustic recording with mobile phones allowed accurate estimation of EF and SV. No previous study or available technology combines mobile phones and acoustic recording in HF diagnosis or monitoring that could be deployed to low-resource settings. The technology represents a novel and potentially far-reaching tool for the screening and management of patients with known and undiagnosed HF.

Acknowledgments

The research team would like to acknowledge research coordinator Alyssa Thomas for working with medical students and residents who enrolled and consented participants for this study. This study was not funded by any agency. As stated in the institutional review board approval, Fleming Scientific served as the sponsor of this research, providing phones and assisting in quality control and analysis of the acoustic data. No funds were transferred among any of the parties in this team research effort. No financial relationships exist among any of the entities involved in this research and study.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to proprietary analytical techniques. Clinical data are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Lam CSP, Solomon SD. Classification of heart failure according to ejection fraction: JACC review topic of the week. J Am Coll Cardiol 2021;77(25):3217-3225 [FREE Full text] [doi: 10.1016/j.jacc.2021.04.070] [Medline: 34167646]
2. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res 2023;118(17):3272-3287 [FREE Full text] [doi: 10.1093/cvr/cvac013] [Medline: 35150240]
3. Psotka MA, Fonarow GC, Allen LA, Maddox KEJ, Fiuzat M, Heidenreich P, et al. The hospital readmissions reduction program: nationwide perspectives and recommendations: a JACC: heart failure position paper. JACC Heart Fail 2020;8(1):1-11 [FREE Full text] [doi: 10.1016/j.jchf.2019.07.012] [Medline: 31606360]
4. Steeds RP. Multimodality imaging in heart failure patients. Curr Opin Cardiol 2013;28(2):209-215. [doi: 10.1097/HCO.0b013e32835bc1b5] [Medline: 23295552]
5. Wood PW, Choy JB, Nanda NC, Becher H. Left ventricular ejection fraction and volumes: it depends on the imaging method. Echocardiography 2014;31(1):87-100 [FREE Full text] [doi: 10.1111/echo.12331] [Medline: 24786629]
6. Klaeboe LG, Edvardsen T. Echocardiographic assessment of left ventricular systolic function. J Echocardiogr 2019;17(1):10-16 [FREE Full text] [doi: 10.1007/s12574-018-0405-5] [Medline: 30390189]
7. Nguyen KL, Hu P, Finn JP. Cardiac magnetic resonance quantification of structure-function relationships in heart failure. Heart Fail Clin 2021;17(1):9-24 [FREE Full text] [doi: 10.1016/j.hfc.2020.08.001] [Medline: 33220890]
8. Bachtiger P, Petri CF, Scott FE, Park SR, Kelshiker MA, Sahemey HK, et al. Point-of-care screening for heart failure with reduced ejection fraction using artificial intelligence during ECG-enabled stethoscope examination in London, UK: a prospective, observational, multicentre study. Lancet Digit Health 2022;4(2):e117-e125 [FREE Full text] [doi: 10.1016/S2589-7500(21)00256-9] [Medline: 34998740]
9. Zile MR, Mehra MR, Ducharme A, Sears SF, Desai AS, Maisel A, et al. Hemodynamically-guided management of heart failure across the ejection fraction spectrum: the GUIDE-HF trial. JACC Heart Fail 2022;10(12):931-944 [FREE Full text] [doi: 10.1016/j.jchf.2022.08.012] [Medline: 36456066]
10. Ramgobin D, Vo M, Golarmari R, Jain R, Jain R. Congestive heart failure clinics and telemedicine: the key to reducing hospital readmissions in the United States. Cardiol J 2022;29(6):1013-1019 [FREE Full text] [doi: 10.5603/CJ.a2021.0073] [Medline: 34308538]
11. Richter F. Charted: there are more mobile phones than people in the world. World Economic Forum. 2023. URL: https://www.weforum.org/agenda/2023/04/charted-there-are-more-phones-than-people-in-the-world/#:~:text=According%20to%20the%20International%20Telecommunication,billion%20halfway%20through%20the%20year [accessed 2024-03-18]
12. Kosmicki DL, Collins SP, Kontos MC, Zuber M, Kipfer P, Jost CA, et al. Noninvasive prediction of left ventricular systolic dysfunction in patients with clinically suspected heart failure using acoustic cardiography. Congest Heart Fail 2010;16(6):249-253 [FREE Full text] [doi: 10.1111/j.1751-7133.2010.00191.x] [Medline: 21091608]
13. Meyer S, Todd D, Wright I, Gortner L, Reynolds G. Review article: non-invasive assessment of cardiac output with portable continuous-wave doppler ultrasound. Emerg Med Australas 2008;20(3):201-208. [doi: 10.1111/j.1742-6723.2008.01078.x] [Medline: 18400002]
14. Radzievsky N, Papyan S, Kushnir I, Gat M, Kushnir A, Sagie A, et al. Estimation of left ventricular function using a novel acoustic-based device. Eur J Clin Invest 2012;42(4):402-410. [doi: 10.1111/j.1365-2362.2011.02596.x] [Medline: 21950619]
15. Salleh SH, Hussain HS, Swee TT, Ting CM, Noor AM, Pipatsart S, et al. Acoustic cardiac signals analysis: a Kalman filter-based approach. Int J Nanomedicine 2012;7:2873-2881 [FREE Full text] [doi: 10.2147/IJN.S32315] [Medline: 22745550]
16. Wang S, Fang F, Liu M, Lam YY, Wang J, Shang Q, et al. Rapid bedside identification of high-risk population in heart failure with reduced ejection fraction by acoustic cardiography. Int J Cardiol 2013;168(3):1881-1886. [doi: 10.1016/j.ijcard.2012.12.064] [Medline: 23352488]
17. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49(11):1225-1231. [doi: 10.1016/s0895-4356(96)00002-9] [Medline: 8892489]
18. Kunyima AB, Lusamba SN, Kunyima MB, Kabele CN. Differential enthalpy, factor of cardiac power and precursor of work power from the nodal tissue. Cardiol Angiol Int J 2016;5(3):1-15 [FREE Full text] [doi: 10.9734/ca/2016/26708]
19. Westerhof N. Cardiac work and efficiency. Cardiovasc Res 2000;48(1):4-7 [FREE Full text] [doi: 10.1016/s0008-6363(00)00176-0] [Medline: 11033101]
20. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49(12):1373-1379 [FREE Full text] [doi: 10.1016/s0895-4356(96)00236-3] [Medline: 8970487]
21. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol 2022;75(1):25-36 [FREE Full text] [doi: 10.4097/kja.21209] [Medline: 35124947]

22. Conway A, Inglis SC, Clark RA. Effective technologies for noninvasive remote monitoring in heart failure. Telemed J E Health 2014;20(6):531-538 [FREE Full text] [doi: 10.1089/tmj.2013.0267] [Medline: 24731212]
23. Iftikhar Z, Lahdenoja O, Tadi MJ, Hurnanen T, Vasankari T, Kiviniemi T, et al. Multiclass classifier based cardiovascular condition detection using smartphone mechanocardiography. Sci Rep 2018;8(1):9344 [FREE Full text] [doi: 10.1038/s41598-018-27683-9] [Medline: 29921933]
24. Indraratna P, Tardo D, Yu J, Delbaere K, Brodie M, Lovell N, et al. Mobile phone technologies in the management of ischemic heart disease, heart failure, and hypertension: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020;8(7):e16695 [FREE Full text] [doi: 10.2196/16695] [Medline: 32628615]
25. MacKinnon GE, Brittain EL. Mobile health technologies in cardiopulmonary disease. Chest 2020;157(3):654-664 [FREE Full text] [doi: 10.1016/j.chest.2019.10.015] [Medline: 31678305]
26. Son YJ, Lee Y, Lee HJ. Effectiveness of mobile phone-based interventions for improving health outcomes in patients with chronic heart failure: a systematic review and meta-analysis. Int J Environ Res Public Health 2020;17(5):1749 [FREE Full text] [doi: 10.3390/ijerph17051749] [Medline: 32156074]
27. Inglis SC, Clark RA, McAlister FA, Stewart S, Cleland JGF. Which components of heart failure programmes are effective? a systematic review and meta-analysis of the outcomes of structured telephone support or telemonitoring as the primary component of chronic heart failure management in 8323 patients: Abridged Cochrane Review. Eur J Heart Fail 2011;13(9):1028-1040 [FREE Full text] [doi: 10.1093/eurjhf/hfr039] [Medline: 21733889]
28. Dierckx R, Pellicori P, Cleland JGF, Clark AL. Telemonitoring in heart failure: big brother watching over you. Heart Fail Rev 2015;20(1):107-116. [doi: 10.1007/s10741-014-9449-4] [Medline: 24972644]
29. Planinc I, Milicic D, Cikes M. Telemonitoring in heart failure management. Card Fail Rev 2020;6:e06 [FREE Full text] [doi: 10.15420/cfr.2019.12] [Medline: 32377385]
30. Winkler S, Schieber M, Lücke S, Heinze P, Schweizer T, Wegertseder D, et al. A new telemonitoring system intended for chronic heart failure patients using mobile telephone technology--feasibility study. Int J Cardiol 2011;153(1):55-58. [doi: 10.1016/j.ijcard.2010.08.038] [Medline: 20851481]
31. Shandhi MMH, Fan J, Heller JA, Etemadi M, Klein L, Inan OT. Estimation of changes in intracardiac hemodynamics using wearable seismocardiography and machine learning in patients with heart failure: a feasibility study. IEEE Trans Biomed Eng 2022;69(8):2443-2455 [FREE Full text] [doi: 10.1109/TBME.2022.3147066] [Medline: 35100106]
32. Stehlik J, Schmalfuss C, Bozkurt B, Nativi-Nicolau J, Wohlfahrt P, Wegerich S, et al. Continuous wearable monitoring analytics predict heart failure hospitalization: the LINK-HF multicenter study. Circ Heart Fail 2020;13(3):e006513 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.119.006513] [Medline: 32093506]
33. Brons M, Koudstaal S, Asselbergs FW. Algorithms used in telemonitoring programmes for patients with chronic heart failure: a systematic review. Eur J Cardiovasc Nurs 2018;17(7):580-588 [FREE Full text] [doi: 10.1177/1474515118786838] [Medline: 29954184]
34. Wang L, Zhou X. Detection of congestive heart failure based on LSTM-based deep network via short-term RR intervals. Sensors (Basel) 2019;19(7):1502 [FREE Full text] [doi: 10.3390/s19071502] [Medline: 30925693]
35. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat Med 2019;25(1):70-74. [doi: 10.1038/s41591-018-0240-2] [Medline: 30617318]
36. Kang SH, Joe B, Yoon Y, Cho GY, Shin I, Suh JW. Cardiac auscultation using smartphones: pilot study. JMIR Mhealth Uhealth 2018;6(2):e49 [FREE Full text] [doi: 10.2196/mhealth.8946] [Medline: 29490899]
37. Dillier R, Kobza R, Erne S, Zuber M, Arand P, Erne P. Noninvasive detection of left-ventricular systolic dysfunction by acoustic cardiography in atrial fibrillation. Cardiol Res Pract 2010;2011:173102 [FREE Full text] [doi: 10.4061/2011/173102] [Medline: 20981304]
38. Wolsk E, Kaye DM, Komtebedde J, Shah SJ, Borlaug BA, Burkhoff D, et al. Determinants and consequences of heart rate and stroke volume response to exercise in patients with heart failure and preserved ejection fraction. Eur J Heart Fail 2021;23(5):754-764 [FREE Full text] [doi: 10.1002/ejhf.2146] [Medline: 33686716]
39. Heironimus J, Stromberg A, Steiner RWP. Detecting COVID-19 respiratory markers with ordinary mobile phones. Int J Digit Healthc 2022;2(1):109 [FREE Full text] [doi: 10.15344/ijdh/2022/109]
40. Policy for device software functions and mobile medical applications: guidance for industry and Food and Drug Administration staff. U.S. Food and Drug Administration. 2022. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications [accessed 2022-12-20]

Abbreviations

AI: artificial intelligence
ASR: Another Sound Recorder
AUROC: area under the receiver operating curve
BARDA: Biomedical Advanced Research and Development Authority
BSA: body surface area

ECG: electrocardiogram
EF: ejection fraction
FDA: US Food and Drug Administration
HF: heart failure
IRB: institutional review board
LVEF: left ventricular ejection fraction
NIH: National Institutes of Health
SV: stroke volume
TSD: Time Series Dynamics

Edited by A Mavragani; submitted 05.02.24; peer-reviewed by T Aslanidis, S Chatterjee, T Dr. Edoh, C Baxter; comments to author 12.03.24; revised version received 19.03.24; accepted 10.04.24; published 26.06.24.

Please cite as:
Huecker M, Schutzman C, French J, El-Kersh K, Ghafghazi S, Desai R, Frick D, Thomas JJ
Accurate Modeling of Ejection Fraction and Stroke Volume With Mobile Phone Auscultation: Prospective Case-Control Study
JMIR Cardio 2024;8:e57111
URL: https://cardio.jmir.org/2024/1/e57111
doi:10.2196/57111
PMID:

Original Paper

Cardiac Rehabilitation During the COVID-19 Pandemic and the Potential for Digital Technology to Support Physical Activity Maintenance: Qualitative Study

Linda G Park[1,2], PhD, FNP-BC; Serena Chi[3], BA; Susan Pitsenbarger[4], BA; Julene K Johnson[5], PhD; Amit J Shah[6], MSCR, MD; Abdelaziz Elnaggar[1,2], MPH, MBBS; Julia von Oppenfeld[2], BA; Evan Cho[7], BS; Arash Harzand[8], MBA, MD; Mary A Whooley[2,9,10], MD

[1]Department of Community Health Systems, University of California San Francisco, San Francisco, CA, United States

[2]Veterans Affairs Medical Center, San Francisco, CA, United States

[3]Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA, United States

[4]Bennett ReGen, Dearborn, MI, United States

[5]Institute for Health & Aging, School of Nursing, University of California San Francisco, San Francisco, CA, United States

[6]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[7]College of Medicine, California Northstate University, Elk Grove, CA, United States

[8]Division of Cardiology, School of Medicine, Emory University, Atlanta, GA, United States

[9]Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[10]Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA, United States

Corresponding Author:
Linda G Park, PhD, FNP-BC
Department of Community Health Systems
University of California San Francisco
2 Koret Way
Room 531A
San Francisco, CA, 94143-0610
United States
Phone: 1 415 502 6616
Email: linda.park@ucsf.edu

Abstract

Background: Social distancing from the COVID-19 pandemic may have decreased engagement in cardiac rehabilitation (CR) and may have had possible consequences on post-CR exercise maintenance. The increased use of technology as an adaptation may benefit post-CR participants via wearables and social media. Thus, we sought to explore the possible relationships of both the pandemic and technology on post-CR exercise maintenance.

Objective: This study aimed to (1) understand CR participation during the COVID-19 pandemic, (2) identify perceived barriers and facilitators to physical activity after CR completion, and (3) assess willingness to use technology and social media to support physical activity needs among older adults with cardiovascular disease.

Methods: We recruited participants aged 55 years and older in 3 different CR programs offered at both public and private hospitals in Northern California. We conducted individual interviews on CR experiences, physical activity, and potential for using technology. We used thematic analysis to synthesize the data.

Results: In total, 22 participants (n=9, 41% female participants; mean age 73, SD 8 years) completed in-depth interviews. Themes from participants' feedback included the following: (1) anxiety and frustration about the wait for CR caused by COVID-19 conditions, (2) positive and safe participant experience once in CR during the pandemic, (3) greater attention needed to patients after completion of CR, (4) notable demand for technology during the pandemic and after completion of CR, and (5) social media networking during the CR program considered valuable if training is provided.

Conclusions: Individuals who completed CR identified shared concerns about continuing physical activity despite having positive experiences during the CR program. There were significant challenges during the pandemic and heightened concerns for safety and health. The idea of providing support by leveraging digital technology (wearable devices and social media for

social support) resonated as a potential solution to help bridge the gap from CR to more independent physical activity. More attention is needed to help individuals experience a tailored and safe transition to home to maintain physical activity among those who complete CR.

(JMIR Cardio 2024;8:e54823) doi:10.2196/54823

KEYWORDS

cardiac rehabilitation; cardiac rehab; COVID-19; digital health; digital technology; physical activity; physical activity maintenance; social media; older adults; pandemic; social distancing; technology; wearables; CR; exercise; cardiovascular disease; gerontology; geriatric; geriatrics; hospital; medical facility; California; interview; thematic analysis; anxiety

Introduction

Cardiac rehabilitation (CR) is a critical aspect of recovery that is offered to adults who experience cardiac events, such as a myocardial infarction, coronary revascularization, and valve replacement, but is significantly underused. CR involves a comprehensive 12-week group program consisting of supervised physical activity training, patient education, and risk modification and is considered an American Heart Association/American College of Cardiology Class IA level recommendation for its health benefits [1]. CR is associated with reduced morbidity and mortality as well as improved quality of life, functional capacity, independence, and symptoms of dyspnea and fatigue [2-6]. Maintaining regular physical activity after CR improves physical function [7,8] and health-related quality of life [4] and is associated with a reduction in the risk of secondary cardiac events, depression, and all-cause mortality [9].

Despite the myriad benefits, many CR participants eventually return to a sedentary lifestyle [10-12] despite the expectation to maintain physical activity independently upon completion. Only 15%-50% report any exercise 6 months after CR completion [10-12], negating the long-term health benefits of CR [13]. Reported barriers to maintaining physical activity after CR include diminished physical condition, competing demands (eg, family health issues), lack of motivation, lack of interest, lack of social support, environmental factors (eg, lack of transportation), and financial costs [14,15]. In addition, participants of CR often receive little to no support during the transition from CR to community- or home-based exercise and desire support mechanisms to ease their transition [16]. The potential enablers to maintaining physical activity after CR include continued contact with CR staff after finishing a CR program, extending the weeks of the CR program, returning for check-ins after CR discharge, having an exercise plan after CR completion, and receiving social support from family and friends [13]. Prior research shows that personal contact is essential to support a successful transition to community-based exercise after CR [16].

In addition to personal contact, the use of personal technology has emerged as a key ally in maintaining an active lifestyle for CR patients. Using technological solutions post phase II CR offers a multitude of benefits for patients. Advanced wearables, such as heart rate monitors and fitness trackers, enable individuals to self-monitor their exercise performance and vital signs, promoting self-awareness and motivation [17]. Mobile apps and telehealth platforms also facilitate access to personalized tracking and provide a digital connection to health care providers [18]. This accessibility enhances patient engagement and adherence to prescribed exercise routines, reducing the risk of relapse and promoting long-term cardiovascular health [19].

The COVID-19 pandemic posed additional challenges related to participation in CR and health-promoting behaviors. A survey study in the United Kingdom reported that COVID-19 lockdown restrictions were associated with significantly decreased participation in CR; changes in CR location, goals, supervision, duration, and enjoyment; and increased perceived effort [20]. A confluence of factors contributed to challenges of participating in CR during the pandemic including programs being forced to suspend or terminate in-person, facility-based services. Declines in CR participation were most marked among dual Medicare and Medicaid enrollees as well as those living in rural areas or socially vulnerable communities based on the Social Vulnerability Index [21]. Compared with CR programs that remained open since the pandemic, the 220 CR centers that closed were more likely to be affiliated with public hospitals located in rural areas and served the most socially vulnerable communities [21].

While mobile phone and social media use has increased among older adults over the past several decades [22], the pandemic showed that older adults rely on and can engage with digital technologies for health care access. Beyond the pandemic, understanding the barriers and experiences of CR participants may lead to more patient-centered CR for those who are unable to or do not wish to participate in facility-based CR and may also be useful for the rapid and effective implementation of CR. This study aimed to (1) describe perceived barriers and facilitators to physical activity after CR completion and (2) understand CR participation during the COVID-19 pandemic among older adults who participated in the ACTION (Americans & Cardiac Rehabilitation Training In Older Adults Needs) study.

Methods

Ethical Considerations

The ACTION study was approved by the institutional review boards from all 3 participating sites (John Muir Medical Center IRC-ID 20-08-02; NorthBay Healthcare NBH 21-05; and University of California, San Francisco IRB 20-31215). Participants provided written informed consent by reading the participant information sheet and signing the participant consent form. All participants were given the opportunity to ask questions. All subject data were deidentified (eg, questionnaires,

recordings, and transcripts were coded with specific ID numbers). Data were encrypted and stored on a password-secured database. We collected both quantitative and qualitative experiences after completing CR and compensated interview participants US $50 and survey participants US $30. The former will be reported in subsequent publications.

Design, Recruitment, and Study Sample

This paper describes a qualitative research study that sought to understand beliefs and experiences related to CR among participants with a history of CR participation. Our approach involved conducting in-depth interviews to obtain open-ended responses that were then organized into themes after generating initial codes (ie, thematic analysis). We also collected quantitative data from numerical ratings on comfort levels with various technologies. Two of the coauthors (SP and SC) conducted interviews and data analysis, with a final agreement on the themes by the principal investigator (LGP).

The recruitment sites included 2 community CR centers and 1 university-affiliated center in Northern California. Recruitment occurred between March and September 2022, a time period when public health directives on mask mandates, strict infection control measures, and social distancing practices were evolving from the COVID-19 pandemic. Inclusion criteria were being 55 years of age or older, between 3 and 24 months post-CR participation (of at least 1 session), having English fluency of moderately well to proficient, and being able to provide informed consent. The exclusion criterion was participation in phase III CR (optional extended CR after outpatient CR for those who pay out-of-pocket; reasons for nonparticipation could be the inability to afford phase III CR or other personal reasons). We used multiple methods to identify possible participants. Initial screening for eligibility and recruitment was done by CR staff who also contacted participants to verify eligibility for study participation. We also screened and recruited participants from a pool of respondents who completed a web-based Qualtrics survey that collected quantitative data related to CR experiences, which will be described in future manuscripts. At the end of the survey, respondents indicated if they were interested in participating in an individual 30-minute phone or video interview. Recruitment continued until data saturation was reached when no new codes arose in the analysis of iterative and open-ended questions.

Data Collection

We collected sociodemographic information such as age, self-identified gender, race, partner status, employment, education, income, and diagnoses for CR. Participants were also asked about their experience with technology. Specifically, they were asked to rate their comfort level with smartphone technology, wearable devices, and social media on a scale of 0-10, with 0 being extremely uncomfortable to 10 being extremely comfortable. The interview guide consisted of 23 questions that were categorized into four major areas: (1) perspectives on their CR experience as a whole, (2) physical activity since completing the CR program, (3) impact of the COVID-19 pandemic on physical activity, and (4) thoughts on various technologies that may aid CR (Multimedia Appendix 1).

Individuals had the choice of interviewing over the phone or video conferencing (ie, Zoom [Zoom Technologies] and Facetime [Apple]); however, all participants elected to be interviewed via phone. All interviews were conducted between June and August 2022 and led by either or both interviewers (SP and SC). All interviews were audio-recorded and transcribed verbatim by a third party. The interviewers also took notes to capture key points and latent data.

Data Analysis

We analyzed interview transcripts using thematic counts [23] to accurately complete an inductive thematic analysis [24]. Each interview transcript underwent a close reading and coding by 2 raters (SP and SC). The data were organized using Microsoft Excel (Microsoft Corp) and then analyzed [25,26]. Upon reviewing the transcript, 2 raters (SP and SC) independently identified key quotes and developed or assigned them to inductive codes. Such independent analysis ensured intercoder reliability and maintained the credibility and dependability of findings [27,28]. After the initial thematic analysis, raters discussed any coding discrepancies until consensus on the final coding scheme and analysis was achieved or settled by the principal investigator. A running count of responses for each code allowed qualitative data to be transformed for a quantitative understanding of patient responses [29]. Sociodemographic and self-reported technology use were summarized using descriptive statistics.

Results

Patient Characteristics and Technology Use

We completed 22 interviews until we achieved data saturation. Table 1 displays the sociodemographic characteristics of the 22 participants, who had a mean age of 73 (SD 8) years, with 41% (n=9) female participants with the majority identifying as White. More than half of the participants were considered low income for living in Northern California. Table 2 outlines the number of participants who were able to complete CR compared to those who had interruptions to CR related to the COVID-19 pandemic.

Table 1. Patient characteristics (N=22).

Characteristics	Values, n (%)
Age (years), mean (SD)	73 (8)
Sex (female)	9 (41)
Race	
Asian	2 (9)
Hispanic	1 (5)
White	19 (86)
With partner	14 (64)
Employed	6 (27)
College graduate	19 (86)
Income ≥US $75,000 per year	10 (71)[a]
Diagnoses for CR[b]	
Ischemic heart disease	17 (77)
Heart failure	2 (9)
Valvular heart disease	3 (14)

[a]A total of 8 participants declined to answer (n=14).

[b]CR: cardiac rehabilitation.

Table 2. Participants enrolled in the CR[a] program between 2020 and 2021.

Timeline (weeks)	Completed CR program (n=16)	CR stopped due to COVID-19 (n=6)
<4	1	0
4-8	3	3
9-12	9	0
≥13	3	1
Unknown	N/A[b]	2

[a]CR: cardiac rehabilitation.

[b]N/A: not available.

The comfort level with technology (smartphone, wearable devices, and social media) was relatively high. Table 2 outlines the number of participants who were able to complete CR compared to those who had interruptions to CR related to the COVID-19 pandemic. Of the 22 participants, 18 owned smartphones, 14 owned a wearable device, and 10 used social media. On a self-reported scale of 0-10 representing the level of comfort using technology, participants reported a mean comfort level of 8.2 with smartphones (n=18), 7.9 with wearable devices (n=14), and 6.8 with social media (n=10).

Overview of Themes

Five themes prevailed from CR participant responses as displayed in Textbox 1 and as detailed as follows. There was congruence in perspectives without distinct differences noted based on CR site or other patient characteristics. One site had additional challenges beyond the pandemic compared to the other sites due to the relocation of CR services from unexpected facility damage.

Textbox 1. Summary of themes from qualitative interviews.

Themes

- Anxiety and frustration about the wait for cardiac rehabilitation caused by COVID-19 conditions
- Positive and safe participant experience once in cardiac rehabilitation during the pandemic
- Greater attention is needed for patients after cardiac rehabilitation completion
- Notable demand for technology during the pandemic and after cardiac rehabilitation completion
- Social media networking during the cardiac rehabilitation program is considered valuable if training is provided

Theme 1: Anxiety and Frustration About the Wait for CR Caused by COVID-19 Conditions

Interviews revealed that participants experienced anxiety and frustration due to the long wait to get into the CR program. The majority of participants (17 of 22) experienced long entry wait periods of between 2 and 3 months to enter the program or faced program closure due to the pandemic. Furthermore, 1 primary on-site location housing the CR program was closed due to facility infrastructure issues, forcing participants to relocate to a more distant location and exacerbating preexisting stressors. In general, delays in getting into the CR program resulted in decreased exercise activity for patients who were newly discharged from the hospital, as they were unable to enter a CR program for many months. Patients expressed their desperation to get into any program, and some called the nearest city with similar programs yet yielded similar wait times. Consequently, patients used resources of personal trainers, internet videos, or were able to get into a home-based physical therapy program. One participant stated:

> *Yeah, it was a long wait. But I was told, when I started the process, it might be three months before I could get in. At one point I made contact with the program in XX County that's affiliated with XX to see if I could get in there. And they also had a long wait list.*

Facing the COVID-19 barrier, patients expressed they felt an urgency to independently create a healthy environment with exercise and nutrition. It was prevalent with all 22 participants that the urgency to get started after discharge from the hospital was strong. However, without direction or exercise monitoring, patient hesitation to start exercising at home was highly noted, as reflected by 1 patient:

> *The thing you always worry about when you exercise after having a surgery like that is am I going to have a heart attack.*

Theme 2: Positive and Safe Participant Experience Once in CR During the Pandemic

Once in the CR program, all of the participants agreed that the program was well organized and professional. After the aggravation of the long wait to enroll in the program for some, patient consensus was that the staff were exceptional in their care and monitoring. One patient explained this:

> *I was very impressed by both the way in which the program operated, and the people themselves in terms of their competence level and their understanding of the issues, and their concern for patients, and their flexibility.*

Many patients considered the program a "safe zone," which increased their level of confidence to exercise. In developing a fitness regime after cardiac surgery, the presence of monitoring was appreciated:

> *...knowing there was a registered nurse for every two people, they monitored you – when you're getting back into physical therapy activity after you've had some kind of heart situation, is comforting.*

One participant described this further:

> *The ability to test all the equipment that had monitors was most effective in giving the patients the freedom to try equipment to see what worked well for them to do the work that needed to be done.*

Overall, all participants had positive comments about the CR program and felt they would have stayed in the program on a long-term basis if their insurance would cover the cost for more sessions.

Theme 3: Greater Attention Needed to Patients After CR Completion

Once patients finished the CR program, many felt on their own and were concerned about keeping up with the regime they learned, as well as lacking the specific equipment provided in the CR program. Voicing concern that there was no follow-up after finishing the program, 1 participant expressed,

> *Maybe they could have a month later call and say to you how are you doing, are you keeping up with your exercises, how many steps are you having a day?*

This quote reflected the sentiment of the majority of the CR program participants. Another said,

> *I would go back and say "can I use those other four appointments?" because I want to see where I am. I know where I am, and it's not where I was when I ended the program.*

Having no one to follow-up on their progress once home, a participant indicated,

> *...(her exercise) is on and off. And so, the only consistent exercise is walking. And so, the weights have dropped off pretty much. The bike work has dropped off pretty much, and I'm having trouble now with my back with being able to walk without pain in it. So, I do walk but I don't walk anywhere near where I'm supposed to be walking as far as the amount of activity I'm supposed to be doing. So, what I planned to do is take my information and start slowly again and build up, pretend I'm at rehab, only it's just me.*

After CR completion, several participants worked and were financially able to hire a personal trainer. Those who worked and did not hire a personal trainer expressed a need to develop greater time management skills to keep up with the needed exercises recommended in the program. Retirees turned to family members and spouses for support with mobile apps, walking together, or going to the community gym.

Theme 4: Notable Demand for Technology During the Pandemic and After CR Completion

While on the waitlist for the CR program during COVID-19, many patients sought technological options (eg, computers, phones, and the internet) to sustain physical activity and self-care. A participant said,

> *Sometimes during Covid, when like everybody was locked down, I did do that on my computer. I did (exercise) classes on my computer.*

Most felt it would be a positive alternative if there was another lockdown for those in a program; additionally, 82% (n=18) of

the participants believed web-based forms to engage in physical activity would be a great option once they finished in the program.

Participants were asked their opinion about the idea of using a wearable device for self-monitoring to help with physical activity after the prescribed CR program (with a prompt related to having a personal coach). The reassurance of a monitor in tracking several health metrics along with personal progress and having a personal coach to review health data were considered most important to individuals who had experienced a life-threatening event. Having these resources would give them a sense of being in a "safe zone." One participant's comment reflected similar responses from the other participants:

> *I'm through the program, but now I have to pay and if there was something like this, you know, with a wrist monitor and you could do your exercises and meet with, you know a personal coach from the rehab, or who's monitoring you, that would be wonderful.*

Finding the resources on "what to do, how to use and what information can I get from this tool" was a key factor in moving forward working with a utility device.

In total, 14 participants stated they owned a wearable device; however, not all participants actively used it. The barrier to use (brought up by 4 participants) was the lack of user friendliness or guidance in learning how to use the device. Some participant quotes were "I can see what I want to see, I can probably learn more," "It's manageable, but I feel my abilities with it are limited," and "I have a smartwatch, but I need help setting it up." When considering using a wearable device for tracking health metrics and monitoring personal progress, a concern arose that "I would need training on how to set it up." A participant considered seeking assistance from family: "my grandchildren would help my wife. They were born with those things, you know?"

Theme 5: Social Media Networking During the CR Program Considered Valuable if Training Was Provided

When asked about their opinion on using social media to supplement their CR experience, half of the participants voiced interest in using social media to see what others were experiencing while undergoing the program or to build camaraderie with other CR participants, especially in the context of limited social interaction during the COVID-19 pandemic. During the height of the pandemic, social distancing precautions prevented CR participants from being able to talk with other attendees and develop friendships with people having similar issues. With reference to having an opportunity to join a Facebook private group with CR participants, someone stated:

> *that would be ideal. Yes. I think this would be ideal because, you know, there's always questions that come up, I think, you know, are you going through this right now or, you know, all this is suddenly affecting me.*

The abilities to support one another and share information were both key components in participants' interest in using Facebook groups. A participant's perspective on its value was to "establish a good rapport with some of these folks in the social media, you become the coach to each other."

Most of the participants were open to trying technology, including Facebook, with the condition there was training or assistance. With social media, some responded they did not know how to maneuver through Facebook beyond keeping up to date with family or browsing posts. A participant shared, "I would do it probably. If I could work the Facebook system to do that, I would." Ultimately, there was agreement on the need for technological options and guidance while in a CR program as well as after finishing the program.

Discussion

Principal Findings

This qualitative study presents the positive experiences participants have in CR, the impact of the COVID-19 pandemic on both physical activity and their CR experience, and the need for additional clinical and social support after CR completion (eg, wearable devices and social media). Along with the 5 themes that have been presented, the findings could also be summarized from the standpoint of stages of CR. More specifically, the way in which technology could be deployed as a solution to barriers at all stages. Before the CR initiation, there was heightened anxiety in getting enrolled into a program due to long pandemic-related waitlists and program closures. During this waiting period, participants expressed fear and hesitation to start exercising without supervision. Technology in the form of positive messages and education regarding safe exercise could be delivered before, during, and after CR to address these concerns. Due to the persistence of long wait times for CR initiation after the pandemic, clinicians may reconsider the role of proper discharge instructions from the hospital and include more details about exercise safety while they wait for phase II CR. During enrollment in CR, there was an increase in participant confidence and comfort with exercise. During and after CR participation, patients reported the potential benefit of wearable health devices to track exercise and to have those data reviewed by their CR team. In addition, patients were enthusiastic about using wearable devices and health apps to support their participation in phase II CR or their health behaviors after completing CR. Furthermore, wearable devices could potentially provide a practical alternative to CR in the case where in-person CR is inaccessible (ie, long enrollment waitlists or another pandemic). After CR, participants expressed concerns about maintaining exercise without equipment and not being able to ask questions to a CR provider. This older group of participants were open to participating in social media networking groups as a means to increase social support and use peers as a resource for answering questions. Some participants cited they could solicit technology support from their family members if needed. These data affirm the benefits of CR, despite pandemic-related barriers, and the positive outlook on using technology as a solution to shortcomings in exercise monitoring and social support.

Although drop-offs in exercise following CR completion are well-known, our study found several reasons that may help design future policy changes or interventions that can increase post-CR exercise. As noted above in theme 3, after program completion, patients had variable access to necessary exercise

equipment and felt a lack of guidance or follow-up from professional staff. Such findings are key for understanding the impacts on long-term physical and mental health. Exercise is critical for maintaining cardiovascular health and reducing the risk of future cardiac events [30], thus policy changes including reimbursement of long-term exercise equipment (eg, stationary bike) or periodic web-based or in-person check-ins with the CR team could be considered. With limited access to exercise equipment or continued training, patients may struggle to maintain their physical health improvements. Additionally, lack of check-ins following CR completion decreases adherence to a healthy lifestyle and increases the likelihood of patients returning to previous unhealthy behaviors such as smoking, unhealthy eating habits, or a sedentary lifestyle, thereby increasing the risk of future cardiac events and negative patient outcomes [31]. In addition to physical effects, the period following CR may have a significant psychological impact as patients may experience depression, anxiety, or fear of future cardiac events [32]. While CR may assist patients in managing these emotions, without continued support, patients may struggle to cope. Evident in our study, patients felt ill-prepared to maintain health improvements and track progress, heavily relying on guesswork with high anxiety. This only magnifies existing mental health risks. Thus, attention to post-CR health remains an area for improvement for health care providers. Further research is needed to determine optimal check-in frequency and methods for ongoing physical and psychological support.

This study specifically addressed the potential role of integrating technology to improve physical activity after CR completion. Traditionally, older adults have been considered resistant to the use of new technologies; however, our study refutes this myth and validates that older adults are eager to engage with certain technologies and have high self-reported comfort levels with using technology. Other recent studies have shown that modern technology, including wearable devices and mobile apps, can be practical tools in maintaining physical activity levels after completing CR [33,34]. In particular, older adults have demonstrated a strong potential for adopting new technologies to support their physical activity and maintain healthy lifestyles [35,36]. For example, in a previous study, we found that a home-based CR program that included wearable activity trackers and web-based support was well received by older adults who reported improved physical activity levels and overall satisfaction with the program [37]. Other studies have found that technology-based interventions can be effective in reducing anxiety and depression in post-CR patients, as well as improving adherence to healthy lifestyle behaviors [38,39]. With the increasing availability and affordability of digital health technologies, there is significant potential to integrate these tools into CR programs to provide ongoing support and improve long-term health outcomes for patients. However, it is crucial to consider older adults' specific needs and preferences in designing and implementing technology-based interventions and provide appropriate training and support to ensure successful adoption and sustained use.

There are several clinical and research implications that can be derived from this study's findings. CR providers should assess participant needs early on in their program to assess psychological status (anxiety or depression) and other barriers (equipment or gym access) to be successful with independent physical activity after CR completion. In addition, providers can assess other needs such as health conditions that require safety precautions for home exercise. Standardized evidence-based practice guidelines are needed to guide participants who graduate from the CR program. Patients need to be taught about long-term habit formation and motivation to continue physical activity after phase II CR is over, as well as discuss a transition plan in advance of finishing the programs with their CR clinical team. In addition, incorporating the use of a wearable device with personal monitoring could alleviate fears of having secondary events as expressed by many participants. Building in the use of digital technology during phase CR II may be helpful so patients have a warm-up period with support for using the same technology independently after CR completion. For research implications, opportunities include collecting and analyzing data on long-term clinical outcomes (rehospitalization and mortality) from diverse populations who receive any versus no support after CR ends. There is also a need for conducting cost analyses on tools such as digital wearable devices and mobile apps to improve health outcomes.

There are also significant policy implications. This study emphasizes the interest of older adults to engage with wearable devices and social media and may be relevant to multiple stakeholders (ie, payors and health systems) in making decisions on what to pay for and how to deploy the technology. Participants' expressed need for additional support after phase II CR to maintain physical activity is also an important policy implication. Extending insurance coverage for maintenance of remote home-based CR services beyond the traditional 12-week program will help participants transition to independent exercise. This covered extension could be in the form of phase III CR or different payment models that fuse remote patient monitoring with coaching for long-term exercise maintenance. Based on this study, the integration of wearable device data into these services may be beneficial, and financial reimbursement and secure implementation remain an area for future investigation.

Limitations

Our sample mostly comprised White individuals and those from higher educational backgrounds; thus, the findings may not be generalizable to other diverse racial groups and those with low educational attainment. We are also unclear whether financial resources were associated with CR attendance among our participants, including ownership of smartphones and wearable technology that may have influenced their opinions about technology use. In addition, we recruited all participants from 3 urban institutions in Northern California; therefore, our sample may not be generalizable to a broader population including rural populations. There may be heterogeneity in the participants' experiences of older adults and their CR experiences with at least 1 completed CR session (versus up to 36 sessions in some programs) within 3-24 months after CR participation. Despite these limitations, this study provides important insights into the lived experiences and perspectives of 22 older adults representing 3 different CR programs. This study confirmed previous research that describes the perceived lack of support

after CR termination [14]. In addition, this study supports the perceived benefits of adding digital technology as a component of providing tailored feedback to CR participants.

Conclusions

Despite the critical role physical activity plays in sustaining patient cardiovascular health improvements, maintaining adequate activity levels after CR proves to be an immense challenge for several reasons. For example, the transition from the supportive and structured environment of CR centers to daily life leaves patients without the guidance and encouragement they once had. This study highlights requests from participants for regular check-ins and support from health professionals, as well as the integration of digital technologies to improve individuals' motivation and accountability in adhering to their exercise routines. Amplified by the COVID-19 pandemic, these challenges demand thoughtful consideration and tailored strategies to ensure sustained adherence to regular physical activity. This study's findings support the opportunity to leverage technology through wearable devices or mobile apps to sustain engagement in healthy lifestyle behavior because they are cost-effective, tailored, and provide motivation and support for patients in the long term. Although these data were collected during a pandemic, the experiences and perspectives of the participants are generalizable in the current environment of CR with the need to support patients after CR completion and the opportunities that technology offers.

Acknowledgments

We are grateful to the health care team at John Muir Cardiac Conditioning, NorthBay Healthcare Cardiac and Pulmonary Rehabilitation, and the University of California, San Francisco Cardiac Rehabilitation and Wellness Center for facilitating this research as well as the participants who provided their valuable input. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health (award P30AG015272). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 15 KB - cardio_v8i1e54823_app1.docx]

References

1. Leon AS, Franklin BA, Costa F, Balady GJ, Berra KA, Stewart KJ, et al. Cardiac rehabilitation and secondary prevention of coronary heart disease: an American Heart Association scientific statement from the Council on Clinical Cardiology (Subcommittee on Exercise, Cardiac Rehabilitation, and Prevention) and the Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity), in collaboration with the American association of Cardiovascular and Pulmonary Rehabilitation. Circulation 2005 Jan 25;111(3):369-376. [doi: 10.1161/01.CIR.0000151788.08740.5C] [Medline: 15668354]
2. Flynn KE, Piña IL, Whellan DJ, Lin L, Blumenthal JA, Ellis SJ, et al. Effects of exercise training on health status in patients with chronic heart failure: HF-ACTION randomized controlled trial. JAMA 2009;301(14):1451-1499 [FREE Full text] [doi: 10.1001/jama.2009.457] [Medline: 19351942]
3. Goel K, Lennon RJ, Tilbury RT, Squires RW, Thomas RJ. Impact of cardiac rehabilitation on mortality and cardiovascular events after percutaneous coronary intervention in the community. Circulation 2011;123(21):2344-2352 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.983536] [Medline: 21576654]
4. Anderson L, Oldridge N, Thompson DR, Zwisler AD, Rees K, Martin N, et al. Exercise-based cardiac rehabilitation for coronary heart disease: cochrane systematic review and meta-analysis. J Am Coll Cardiol 2016;67(1):1-12 [FREE Full text] [doi: 10.1016/j.jacc.2015.10.044] [Medline: 26764059]
5. Dibben GO, Dalal HM, Taylor RS, Doherty P, Tang LH, Hillsdon M. Cardiac rehabilitation and physical activity: systematic review and meta-analysis. Heart 2018;104(17):1394-1402 [FREE Full text] [doi: 10.1136/heartjnl-2017-312832] [Medline: 29654095]
6. de Araújo Pio CS, Marzolini S, Pakosh M, Grace SL. Effect of cardiac rehabilitation dose on mortality and morbidity: a systematic review and meta-regression analysis. Mayo Clin Proc 2017;92(11):1644-1659. [doi: 10.1016/j.mayocp.2017.07.019] [Medline: 29101934]
7. Forman DE, Arena R, Boxer R, Dolansky MA, Eng JJ, Fleg JL, et al. Prioritizing functional capacity as a principal end point for therapies oriented to older adults with cardiovascular disease: a scientific statement for healthcare professionals

from the American Heart Association. Circulation 2017;135(16):e894-e918 [FREE Full text] [doi: 10.1161/CIR.0000000000000483] [Medline: 28336790]
8. 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;71(19):2149-2161 [FREE Full text] [doi: 10.1016/j.jacc.2018.03.022] [Medline: 29747836]
9. Barons MJ, Turner S, Parsons N, Griffiths F, Bethell H, Weich S, et al. Fitness predicts long-term survival after a cardiovascular event: a prospective cohort study. BMJ Open 2015;5(10):e007772 [FREE Full text] [doi: 10.1136/bmjopen-2015-007772] [Medline: 26493455]
10. Hellman EA. Use of the stages of change in exercise adherence model among older adults with a cardiac diagnosis. J Cardiopulm Rehabil 1997;17(3):145-155. [doi: 10.1097/00008483-199705000-00001] [Medline: 9187980]
11. Moore SM, Charvat JM, Gordon NH, Pashkow F, Ribisl P, Roberts BL, et al. Effects of a CHANGE intervention to increase exercise maintenance following cardiac events. Ann Behav Med 2006;31(1):53-62. [doi: 10.1207/s15324796abm3101_9] [Medline: 16472039]
12. Moore SM, Ruland CM, Pashkow FJ, Blackburn GG. Women's patterns of exercise following cardiac rehabilitation. Nurs Res 1998;47(6):318-324. [doi: 10.1097/00006199-199811000-00005] [Medline: 9835487]
13. Bock BC, Carmona-Barros RE, Esler JL, Tilkemeier PL. Program participation and physical activity maintenance after cardiac rehabilitation. Behav Modif 2003;27(1):37-53. [doi: 10.1177/0145445502238692] [Medline: 12587259]
14. Fleury J, Lee SM, Matteson B, Belyea M. Barriers to physical activity maintenance after cardiac rehabilitation. J Cardiopulm Rehabil 2004;24(5):296-305; quiz 306. [doi: 10.1097/00008483-200409000-00002] [Medline: 15602147]
15. Neubeck L, Freedman SB, Clark AM, Briffa T, Bauman A, Redfern J. Effects of uninephrectomy and high protein feeding in cyclosporine nephropathy. Eur J Prev Cardiol 2012;19(3):494-503 [FREE Full text] [doi: 10.1038/ki.1989.179]
16. Clark AM, Mundy C, Catto S, MacIntyre PD. Participation in community-based exercise maintenance programs after completion of hospital-based cardiac rehabilitation: a mixed-method study. J Cardiopulm Rehabil Prev 2011;31(1):42-46. [doi: 10.1097/HCR.0b013e3181f68aa6] [Medline: 21037483]
17. Kang HS, Exworthy M. Wearing the future-wearables to empower users to take greater responsibility for their health and care: scoping review. JMIR Mhealth Uhealth 2022;10(7):e35684 [FREE Full text] [doi: 10.2196/35684] [Medline: 35830222]
18. Bhaskar S, Bradley S, Chattu VK, Adisesh A, Nurtazina A, Kyrykbayeva S, et al. Telemedicine as the new outpatient clinic gone digital: position paper from the pandemic health system resilience program (REPROGRAM) international consortium (Part 2). Front Public Health 2020;8:410 [FREE Full text] [doi: 10.3389/fpubh.2020.00410] [Medline: 33014958]
19. Yang Z, Jia X, Li J, Mei Z, Yang L, Yan C, et al. Efficacy and safety of Hybrid Comprehensive Telerehabilitation (HCTR) for cardiac rehabilitation in patients with cardiovascular disease: a systematic review and meta-analysis of randomized controlled trials. Occup Ther Int 2023;2023:5147805 [FREE Full text] [doi: 10.1155/2023/5147805] [Medline: 37593110]
20. Kirwan R, Perez de Heredia F, McCullough D, Butler T, Davies IG. Impact of COVID-19 lockdown restrictions on cardiac rehabilitation participation and behaviours in the United Kingdom. BMC Sports Sci Med Rehabil 2022;14(1):67 [FREE Full text] [doi: 10.1186/s13102-022-00459-5] [Medline: 35418304]
21. Varghese MS, Beatty AL, Song Y, Xu J, Sperling LS, Fonarow GC, et al. Cardiac rehabilitation and the COVID-19 pandemic: persistent declines in cardiac rehabilitation participation and access among US medicare beneficiaries. Circ Cardiovasc Qual Outcomes 2022;15(12):e009618 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.122.009618] [Medline: 36314139]
22. Mobile fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2023-05-20]
23. Maxwell JA. Using numbers in qualitative research. Qual Inq 2010;16(6):475-482. [doi: 10.1177/1077800410364740]
24. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1609406917733847 [FREE Full text] [doi: 10.1177/1609406917733847]
25. Bree RT, Gallagher G. Using microsoft excel to code and thematically analyze qualitative data: a simple, cost-effective approach. AISHE-J 2016;6(2):2811-2825.
26. Ose SO. Using excel and word to structure qualitative data. J Appl Soc Sci 2016;10(2):147-162. [doi: 10.1177/1936724416664948]
27. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020;19:160940691989922 [FREE Full text] [doi: 10.1177/1609406919899220]
28. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006;27(2):237-246. [doi: 10.1177/1098214005283748]
29. Creswell JW. Choosing a mixed methods design. In: Clark VLP, editor. Designing and Conducting Mixed Methods Research. Thousand Oaks, CA: SAGE Publications; 2006:58-88.
30. Oldridge N. Exercise-based cardiac rehabilitation in patients with coronary heart disease: meta-analysis outcomes revisited. Future Cardiol 2012;8(5):729-751. [doi: 10.2217/fca.12.34] [Medline: 23013125]
31. Beatty AL, Li S, Thomas L, Amsterdam EA, Alexander KP, Whooley MA. Trends in referral to cardiac rehabilitation after myocardial infarction: data from the national cardiovascular data registry 2007 to 2012. J Am Coll Cardiol 2014;63(23):2582-2583 [FREE Full text] [doi: 10.1016/j.jacc.2014.03.030] [Medline: 24768872]
32. Leung YW, Flora DB, Gravely S, Irvine J. Carving out time for cardiac rehabilitation: a randomized controlled trial. Eur J Prev Cardiol 2018;25(4):408-417.

33. Bentley CL, Powell L, Potter S, Parker J, Mountain GA, Bartlett YK, et al. The use of a smartphone app and an activity tracker to promote physical activity in the management of chronic obstructive pulmonary disease: randomized controlled feasibility study. JMIR Mhealth Uhealth 2020;8(6):e16203 [FREE Full text] [doi: 10.2196/16203] [Medline: 32490838]
34. Pope ZC, Barr-Anderson DJ, Lewis BA, Pereira MA, Gao Z. Use of wearable technology and social media to improve physical activity and dietary behaviors among college students: a 12-week randomized pilot study. Int J Environ Res Public Health 2019;16(19):3579 [FREE Full text] [doi: 10.3390/ijerph16193579] [Medline: 31557812]
35. Chen W, Zhang Z, Giordani B, Larson JL. Technology-enhanced 4active intervention impacting psychological well-being and physical activity among older adults: a pilot study. Int J Environ Res Public Health 2022;19(1):556 [FREE Full text] [doi: 10.3390/ijerph19010556] [Medline: 35010813]
36. Hosteng KR, Simmering JE, Polgreen LA, Cremer JF, Segre AM, Francis SL, et al. Multilevel mHealth intervention increases physical activity of older adults living in retirement community. J Phys Act Health 2021;18(7):851-857 [FREE Full text] [doi: 10.1123/jpah.2020-0592] [Medline: 34039774]
37. Elnaggar A, von Oppenfeld J, Whooley MA, Merek S, Park LG. Applying mobile technology to sustain physical activity after completion of cardiac rehabilitation: acceptability study. JMIR Hum Factors 2021;8(3):e25356 [FREE Full text] [doi: 10.2196/25356] [Medline: 34473064]
38. Dale LP, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R. Text message and internet support for coronary heart disease self-management: results from the Text4Heart randomized controlled trial. J Med Internet Res 2015;17(10):e237 [FREE Full text] [doi: 10.2196/jmir.4944] [Medline: 26490012]
39. Wells A, Reeves D, Capobianco L, Heal C, Davies L, Heagerty A, et al. Improving the effectiveness of psychological interventions for depression and anxiety in cardiac rehabilitation: PATHWAY-a single-blind, parallel, randomized, controlled trial of group metacognitive therapy. Circulation 2021;144(1):23-33 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.120.052428] [Medline: 34148379]

Abbreviations

ACTION: Americans & Cardiac Rehabilitation Training In Older Adults Needs
CR: cardiac rehabilitation

Edited by A Mavragani; submitted 23.11.23; peer-reviewed by H Aghayan Golkashani, M Pritchard; comments to author 16.01.24; revised version received 23.01.24; accepted 24.01.24; published 14.03.24.

Please cite as:
Park LG, Chi S, Pitsenbarger S, Johnson JK, Shah AJ, Elnaggar A, von Oppenfeld J, Cho E, Harzand A, Whooley MA
Cardiac Rehabilitation During the COVID-19 Pandemic and the Potential for Digital Technology to Support Physical Activity Maintenance: Qualitative Study
JMIR Cardio 2024;8:e54823
URL: https://cardio.jmir.org/2024/1/e54823
doi:10.2196/54823
PMID:38483450

Review

Persuasive Systems Design Trends in Coronary Heart Disease Management: Scoping Review of Randomized Controlled Trials

Eunice Eno Yaa Frimponmaa Agyei[1], PhD; Akon Ekpezu[1], MSc; Harri Oinas-Kukkonen[1], PhD

Oulu Advanced Research on Service and Information Systems, Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland

Corresponding Author:
Eunice Eno Yaa Frimponmaa Agyei, PhD
Oulu Advanced Research on Service and Information Systems
Faculty of Information Technology and Electrical Engineering
University of Oulu
Pentti Kaiteran katu 1
Oulu, 90570
Finland
Phone: 358 0449511559
Email: eunice.agyei@oulu.fi

Abstract

Background: Behavior change support systems (BCSSs) have the potential to help people maintain healthy lifestyles and aid in the self-management of coronary heart disease (CHD). The Persuasive Systems Design (PSD) model is a framework for designing and evaluating systems designed to support lifestyle modifications and health behavior change using information and communication technology. However, evidence for the underlying design principles behind BCSSs for CHD has not been extensively reported in the literature.

Objective: This scoping review aims to identify existing health BCSSs for CHD, report the characteristics of these systems, and describe the persuasion context and persuasive design principles of these systems based on the PSD framework.

Methods: Using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines, 3 digital databases (Scopus, Web of Science, and MEDLINE) were searched between 2010 to 2022. The major inclusion criteria for studies were in accordance with the PICO (Population, Intervention, Comparison, and Outcome) approach.

Results: Searches conducted in the databases identified 1195 papers, among which 30 were identified as eligible for the review. The most interesting characteristics of the BCSSs were the predominant use of primary task support principles, followed by dialogue support and credibility support and the sparing use of social support principles. Theories of behavior change such as the Social Cognitive Theory and Self-Efficacy Theory were used often to underpin these systems. However, significant trends in the use of persuasive system features on par with behavior change theories could not be established from the reviewed studies. This points to the fact that there is still no theoretical consensus on how best to design interventions to promote behavior change in patients with CHD.

Conclusions: Our results highlight key software features for designing BCSSs for the prevention and management of CHD. We encourage designers of behavior change interventions to evaluate the techniques that contributed to the success of the intervention. Future research should focus on evaluating the effectiveness of the interventions, persuasive design principles, and behavior change theories using research methodologies such as meta-analysis.

(JMIR Cardio 2024;8:e49515) doi:10.2196/49515

KEYWORDS

coronary heart disease; persuasive systems design; behavior change; randomized controlled trial; RCT; controlled trials; heart; CHD; cardiovascular

Introduction

Coronary heart disease (CHD), also referred to as coronary artery disease (CAD), is the third leading cause of death worldwide and is associated with 17.8 million deaths annually [1]. Despite its significant association with a high mortality rate, it is preventable. With risk factors such as a sedentary lifestyle, physical inactivity, smoking, poor diet, hypertension, and obesity, both pharmacological and nonpharmacological interventions have been proposed to mitigate this menace [2]. Existing evidence suggests that preventing CHD requires lifestyle and health behavior changes [3]. Digital interventions, particularly behavior change support systems (BCSSs), have the potential to reduce risky health behaviors, improve the well-being of the user, and promote healthy lifestyles in patients with CHD. These are information systems that are designed to form, alter, or reinforce the attitudes, behaviors, or compliance of their users voluntarily [4]. A key element in behavior and attitude change is persuasion; the intention to change the behavior of an individual via persuasion may lead to a positive behavioral outcome [5]. Over the past decade, BCSSs and persuasive design have received elaborate attention with strategies that stem from behavior change theories. Although the development of these systems has increased at a startling rate to promote behavior change, the persuasion context (ie, the interdependencies between the user, technology, and the problem domain) and persuasive systems design principles are often ignored [6].

The existing literature reviews have predominantly focused on determining the effectiveness of different kinds of health BCSSs in changing lifestyle behavior, controlling modifiable risk factors, and improving CHD patient outcomes using mobile technologies [7,8], web-based technologies [9], and telerehabilitation [10]. However, evidence on how these systems were developed (ie, the design principles) to achieve the reported behavior change outcomes is not clear [11,12]. Moreover, these studies have focused on specific technologies; hence, the evidence cannot be generalized for all CHD BCSSs. Designing for behavior change involves identifying behavioral goals [5] and gaining an understanding of the behavior change context including the behavior change strategies, system features, and theoretical foundations [13,14] that underpin it.

This scoping review seeks to address this gap by providing an overview of persuasive context and behavior change strategies that support the management of CHD. Identifying these features will provide designers and researchers with an understanding of the persuasion context and persuasive features in systems that seek to promote behavior change in patients with CHD. More specifically, this review seeks to answer the broad review question, What persuasive systems design trends are evident in the management of CHD? To answer this question, this review aims to identify existing health BCSSs, report the characteristics of these systems, and describe the persuasion context and persuasive system design principles of the identified BCSSs for CHD using the Persuasive Systems Design (PSD) model proposed by Oinas-Kukkonen and Harjumaa [13].

The PSD model is the most-used framework for designing and evaluating persuasive systems [15]. Built on theories from psychology, information systems, and other disciplines, the PSD model guides the analysis of the persuasion context, including recognizing the intent of the persuasion, understanding the persuasion event, and defining the strategies in use [13]. Recognizing the persuasion intent involves understanding the roles of the persuader, the persuadee, the change type (ie, compliance, behavior, and attitude change), and the outcome (ie, forming, altering, or reinforcing compliance, behavior, and attitude change). Understanding the persuasion event entails features and characteristics arising from the problem domain (ie, use context), the user (ie, user context), and technology (ie, technology context). Persuasive systems have information content (ie, message) and software features. Defining the strategy involves crafting the content of the message to be delivered, deciding how to present arguments, the route to deliver the message, and the persuasiveness of the message. The model also provides a structure for designing and evaluating persuasive systems based on 7 key postulates and 4 system design principles, including primary task support (supports the user to carry out the actions that will lead to the desired behavior), dialogue support (facilitates the interaction between the user and the technology), credibility support (enhances the perception of trust and reliability of the system), and social support (aids behavior change by leveraging social influence). Table 1 describes specific variables of the PSD model used in evaluating and analyzing the persuasion context and system design features of the reviewed studies.

Table 1. Description of Persuasive Systems Design (PSD) variables used in the analysis, based on Oinas-Kukkonen and Harjumaa [13].

Factor	Description
Analyzing the persuasion context	
The intent	
Intended outcome/change	The intended outcomes of interest in this review will be classified as clinical outcomes, behavioral outcomes, psychological outcomes, and improved quality of life.
Designer/persuader bias	This refers to the unintentional influence that designers, developers, or creators have on the features, content, or functionalities of an intervention due to their viewpoints, experiences, or preferences.
The event	
Use context	The general application domain is health—specifically, CHD[a] as a preventable health condition.
User context	The population of interest will be patients living with CHD.
Technology context	This will include any technological platforms including wearable devices, mobile apps, and web apps.
The strategy	
Message	This will describe the content delivered to inform or educate the user to change their behavior.
Route	This will describe how information and content are delivered to users via the direct route, which uses the user's cognition, or the indirect route using societal cues.
Persuasive principles	
Primary task support	System features that support CHD to perform their primary task by reducing the cognitive load associated with the activity
Dialogue support	System features that provide computer-human dialogue as a means of reinforcing and motivating patients with CHD to perform the primary task
Credibility support	System features that make the patients with CHD believe that the intervention is reliable and credible
Social support	System features that motivate patients with CHD using social influence

[a]CHD: coronary heart disease.

Drawing upon the PSD model, this research aims to provide an overview of BCSSs for managing CHD, report the characteristics of these systems, and describe the persuasion context and persuasive design principles. The rest of the paper is organized as follows: The Methods section describe how this research was conducted. The Results section reveals the findings of the research. This is followed by the Discussion section, which includes the implications of the findings and the conclusion.

Methods

Identification of Studies

This scoping review was conducted by following PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [16]. We conducted a scoping literature search in the Scopus, Web of Science, and MEDLINE electronic databases from 2010 to 2022 using the Population, Intervention, Comparison, and Outcome (PICO) approach. The search domains included patients with CHD aged 18 years and older (population), randomized controlled trial (RCT) and BCSSs for CHD (intervention), and behavior change (outcome). The following search string related to the PICO approach was used ("coronary artery disease" OR "coronary heart disease" OR "ischemic heart disease") AND ("mobile" OR "smart phone" OR "smartphone" OR "web" OR "internet") AND ("prevention" OR "intervention"). The literature search was limited to studies published from 2010 because the PSD framework was proposed in 2009. Thus, including studies from 2010 would reveal evidence-based research trends. The first and second authors (EEYFA and AE, respectively) carried out the literature search and study selection independently. Divergent opinions on study inclusion were resolved through consensus among the 3 authors.

Data Inclusion and Exclusion Criteria

The titles and abstracts were screened for keywords by the first and second authors. These authors downloaded the full text and examined if they were suitable using the following criteria. First, the study had to be an RCT on CHD, published in a peer-reviewed academic journal or conference, and written in the English language. Additionally, the study intervention had to be technology-mediated (ie, used mobile, web, or internet-based applications). Finally, the study intervention had to have the aim of promoting behavioral change, such as physical activity, diet, or smoking cessation to manage or prevent CHD.

Data Extraction

Using the inclusion and exclusion criteria previously outlined, 30 papers were selected and reviewed. These articles were reviewed using the PSD framework. Data extraction and coding were conducted by the first and second authors independently. The 2 authors read the articles, identified the textual descriptions applicable to the design features, and coded them in a Microsoft Excel (Microsoft Corp) spreadsheet. The third author (HO-K) verified and validated the extracted data. Disagreements were

resolved by revisiting the specific papers and reviewing them together until a consensus was reached.

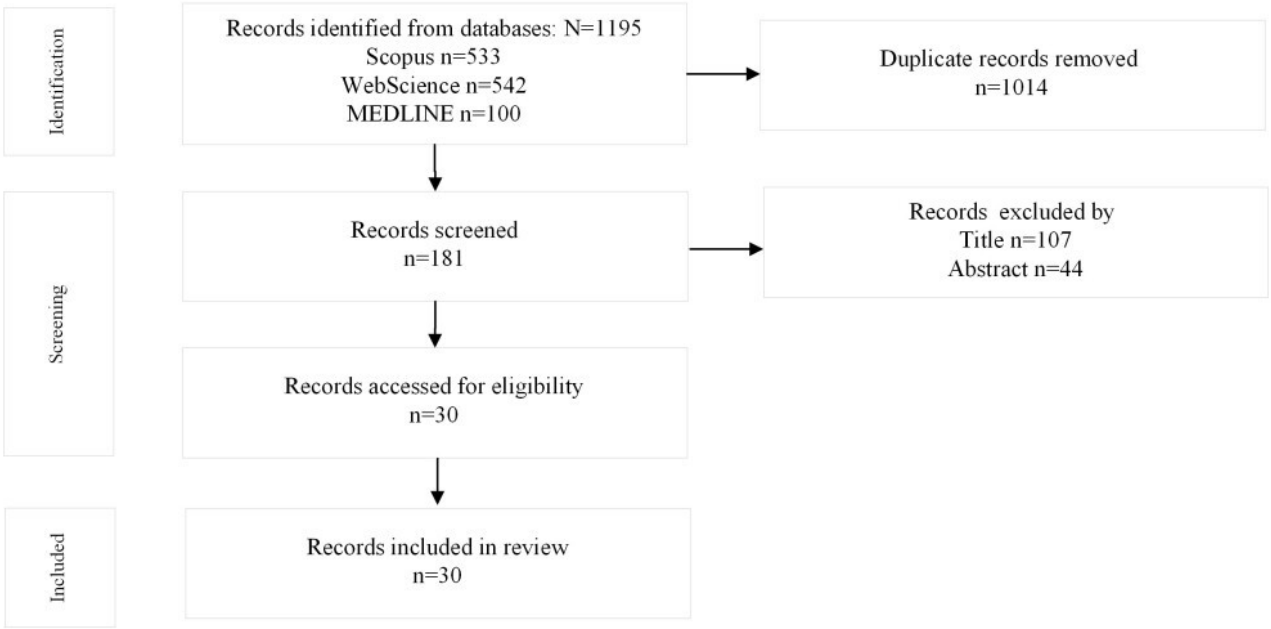
The characteristics of the persuasion context and persuasive design features were extracted. The data extracted from each article were as follows: (1) name of the intervention; (2) objective of the study, which reveals the intention of the intervention; (3) primary and secondary outcome(s) (ie, intended outcome); (4) features/characteristics of the problem domain (ie, use context); (5) description of the study participants (ie, user context); (6) description of the technology-dependent characteristics of the intervention (ie, technology context); (7) description of the target behavior, and (8) persuasive software features of the intervention.

Results

Selection of Studies

The initial search produced 1195 articles that were distributed among the aforementioned databases. Duplicates were removed, leaving 181 papers. The articles were excluded by title if they did not contain the keywords “CHD” and “randomized controlled trial, ” while articles were excluded by abstract if the technological context was not mentioned and if the intervention was not targeted at either behavior change or clinical outcome. Consequently, 151 articles were excluded by title and abstract. Finally, 30 RCT studies were analyzed in this scoping review. Figure 1 highlights the selection process using the PRISMA-ScR flow diagram.

Figure 1. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flow diagram of the study selection process.



Characteristics of the Included Studies

The 30 included RCTs were published in 16 different peer-reviewed journals, namely *Journal of Medical Internet Research* (n=5, 17%), *JMIR MHealth and Uhealth* (n=4, 13%), *European Journal of Preventive Cardiology* (n=3, 13%), *Journal of Cardiopulmonary Rehabilitation and Prevention* (n=3, 10%), *Hearts* (n=2, 7%), *BMJ Open* (n=2, 7%), *Circulation* (n=2, 7%), and the rest in *JAMA*, *Lancet Digital Health*, *Plos ONE*, *Patient Education and Counseling*, *Journal of Cardiovascular Translational Research*, *Pharmacy Education*, *JAMA Cardiology*, *Coronary Artery Disease*, and *Health and Quality of Life Outcomes*.

The study duration for all included RCTs ranged between 4 weeks to 52 weeks, and the sample size for study participants ranged between 84 and 879. Additionally, 70% (n=21) of the studies focused on patients with CHD, 10% (n=3) on patients with acute coronary syndrome, 7% (n=2) on patients with CAD, and the remaining 13% (n=4) on patients with CHD and diabetes, CHD and depression, ischemic heart disease, and clinical manifestation of atherosclerosis in the coronary, cerebral, or peripheral arteries.

Analyzing the Persuasion Context

Overview

As highlighted in Table 1, analyzing the persuasion context involves recognizing the intent of the persuasion, understanding the persuasion event, and defining the strategies in use.

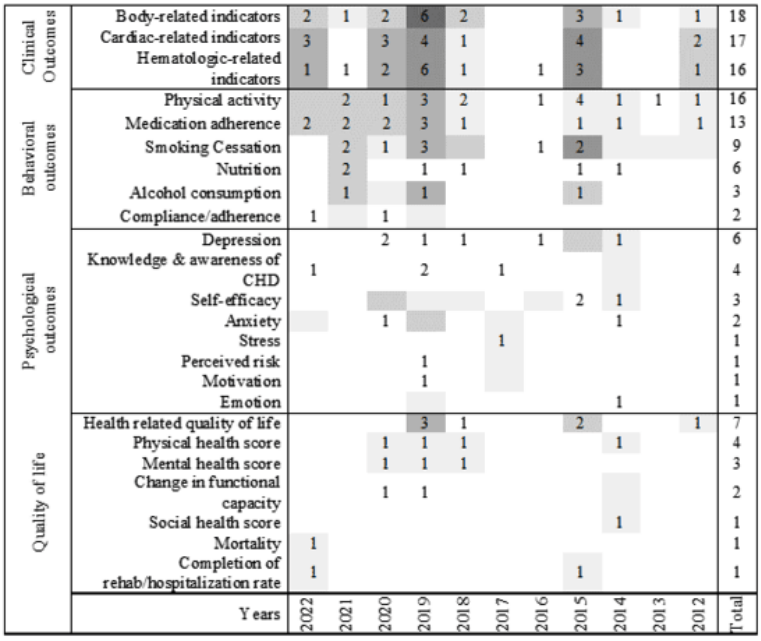
The Intent

All studies stated the objective of their study, which reveals the intention of the persuader as well as the intention behind the intervention. The intention behind the interventions was to use information and communication technology to support patients with CHD to improve their health and lifestyle. The persuaders were researchers (ie, people who participated in the design), and in some cases, clinicians (when the study was intended to be used in a clinical setting). The persuadees were study participants who received the intervention in the RCT. In the design of the interventions, decisions regarding features and content were mostly influenced by the viewpoints, experiences,

and assumptions of the design team. Only 17% (5/30) of studies involved users directly. These users were primarily involved in creating messages delivered through the intervention [17-21]. One (3%) study obtained feedback from users via a pilot test to improve the intervention [22]. Two (7%) studies disclosed the composition of their design team, yet users were not involved in the process [23,24]. Finally, 73% (n=22) of studies did not provide any information on the users/prospective users' involvement in the design of the intervention. Given the potential effect of designer bias on the usability and effectiveness of an intervention, it is important to acknowledge this and put in measures to mitigate its effects.

We identified and classified the intended outcomes of the 30 included RCTs into 4 categories: clinical outcomes, behavioral outcomes, psychological outcomes, and improved quality of life. The distribution of identified intended outcomes in each publication year of the included studies is presented in Figure 2. Multimedia Appendix 1 contains more detailed information [17-46].

Figure 2. Distribution of intended outcomes. CHD: coronary heart disease.



The Event

Analyzing the event provided insights into the problem domain (use context), user characteristics (user context), and technology (technology context). Although the main user group was people who had been diagnosed with CHD, some of them had multiple health conditions, such as diabetes and depression, or needed to change at least one behavior (eg, smoking, physical inactivity, and medication adherence) to prevent further health complications. The varying characteristics of the use context presented unique scenarios and opportunities for health behavior change (Multimedia Appendix 2 [17-46]). Users were mainly older than 18 years. A total of 20 (67%) studies [17-22,24,25,28,30,31,33,35-37,39,40,43,46] required the users to be able to read and understand the text, while 2 studies (7%) [18,29] required users to be computer literate.

The RCTs were conducted on various continents, including Oceania, Asia, Europe, and North America. In Oceania, studies were conducted in Australia [17,20,30,43,45,46] and New Zealand [26,33-35,38]. In Asia, there were 6 (20%) studies conducted in China [19,21,27,28,32,36,42], 3% (n=1) from Singapore [23], 3% (n=1) from the Republic of Korea [25], 3% (n=1) from Indonesia [39], and 3% (n=1) from Pakistan [31]. In Europe, 7% (n=2) of studies were conducted in the United Kingdom [18,40], 3% (n=1) in Italy [44], 3% (n=1) in the Netherlands [22], 3% (n=1) in Belgium [29], and 3% (n=1) across multiple European countries [41]. Two (7%) studies were conducted in North America, namely, the United States [37] and Canada [24].

Furthermore, it was observed that the nationality and cultural factors of users influenced the design of the intervention. For example, in Dorje et al [27], the popular Chinese social media app WeChat, together with Chinese avatars, was used to educate patients. Moreover, different types of technologies were chosen for the design and development for various reasons. The selection of the type of technology was based on some characteristics of the problem and user domain. For example, an SMS or other type of text message was used for delivering the intervention in one study because of its ease of use and cost-effectiveness [20]. In China, WeChat, which has an instant messaging component, was used in 3 (10%) studies [27,36,42].

The Strategy

Exactly 25 (83%) studies stated that they used content in their intervention. The content presented to the users in the 25 studies was mainly to educate users and support the behavior change process. This implies that the user's cognition was required to process the information presented to them. Additionally, the RCTs conducted in Oceania and North America presented content to their users via the interventions, while 2 (7%) RCTs in both Asia and Europe did not present any content. Multimedia Appendices 3 and 4 [17-46] contain additional details.

Within the included RCTs, half (n=15, 50%) incorporated behavior change theories in their intervention designs. The most used behavior change theory was Social Cognitive Theory (n=10, 33%) [17,19-21,24,28,31,35,38,45]. This was followed by Self-Efficacy Theory, which was used in 17% (n=5) of studies [26,33,34,37,38]. Two (7%) studies [17,20], used the Behavior Skill Model, Theory of Reasoned Action, Theory of Planned Behavior, and Control Theory. The Health Belief Model was used in 2 studies [23,31] and the Commonsense Model was used in 2 studies [26,35]. Meanwhile, the Health Action Process Approach theory [28] and Cognition and Behavior theory [42] were each used once. It was observed that the number of theories used across these studies varied from 1 to 5. Interestingly, none of the RCTs conducted in Europe used behavior change theories in their interventions (Multimedia Appendix 4).

Analyzing the Persuasive Features

Overview

System features that were used in the CHD interventions were identified and coded based on the PSD model software feature categories. As shown in Table 1, they comprised primary task support, dialogue support, credibility support, and social support. Multimedia Appendix 4 highlights the distribution of identified persuasive system features. Apart from Putra et al [39], all the analyzed interventions used a minimum of 1 persuasive feature and a maximum of 6 persuasive features. Features of the primary task support principle were the most used, while that of the social support principle was the least used in the interventions for patients with CHD (Multimedia Appendix 5 [17-46]).

Primary Task Support

Primary task support features assist the user in carrying out the primary tasks that lead to behavior change; they include personalization, self-monitoring, reduction, rehearsals, tunneling, tailoring, and simulation [13]. Personalization (n=23, 77%) and self-monitoring (n=18, 60%) were found to be the most widely represented primary task support features in the articles reviewed. Tailoring was identified in 6 (20%) studies. The reduction feature was identified in 1 (3%) study to grade tasks for users [18]. Features such as rehearsal, tunneling, and simulation were not identified. Personalization and semipersonalization were implemented in different forms, such as (1) selecting and providing educational content on CHD based on user characteristics, such as age and gender [40], and baseline characteristics [17,20,43], such as smoking status and diet; (2) using the preferred name of users [19,21,26,35,37] (semipersonalization); (3) using their smoking status and diet pattern [25] (semipersonalization); (4) sending messages at the preferred time of the user [26,35]; (5) individualized exercise programs [27,29,33,38,42]; (6) individualized feedback [28,45]; (7) customizable sounds for reminders [46]; (8) a personalized website based on risk factors [22]; and (9) personalized medication SMS text messages [31].

Self-monitoring was implemented by setting and tracking goals [18,24,40], monitoring physical activities [26,27,29,30,34,44], glucose monitoring [19], heart rate monitoring [27,42], monitoring of cardiovascular risk factors [45], blood pressure monitoring [21,32], step counting with a pedometer [26], and medication adherence using electronic pill bottles [37]. Different messages were tailored for different user groups. For example, different messages were delivered to smokers and nonsmokers [17,20,35,43], dietary messages for vegetarians and nonvegetarians [25] and tailored programs for the working population [23].

Dialogue Support

This support category comprises 7 system features, namely reminders, praise, rewards, liking, similarity, suggestions, and social role. Dialogue support incorporates forms of social or interpersonal interactions into feedback to encourage the user to respond to requests made by the intervention that may lead to behavior change [47]. Reminders, praise, rewards, suggestions, and social role were identified in the studies.

Reminders were the most used software feature in this category (n=13, 43%). Some examples of how reminders were implemented include reminders about behavior change to decrease CHD risk [17,20,43], exercise reminders [23,38], reminders to check blood pressure [27], push notifications [30], medication reminders [31,32,36,37,46], and checkup reminders [26]. This was followed by praise (n=12, 40%). The implementation of praise included individualized feedback to manage outcomes [27,28], motivational messages [29,33,41], personalized feedback on progress [33,40,42,45], and performance [18,26,34].

In addition to these, suggestions (n=3, 10%), rewards (n=10, 3%), and social roles (n=10, 3%) were used, albeit rarely. Suggestions were implemented by providing tips on overcoming hindrances [40], dietary recommendations [46], and information on various physical activities [33]. The social role feature was implemented via a virtual cardiologist coach who offered advice to users [27]. Reward in the form of social rewards was identified in the intervention by Devi et al [18]. Though the liking and similarity feature may have been present in the interventions, we could not evaluate it due to its subjective nature.

System Credibility Support

Design principles in this category support the credibility of the system as a function of persuasion. Its features include expertise, surface credibility, authority, third-party endorsement, real-world feel, trustworthiness, and verifiability. Here, the visual elements of the user interface, as well as the believability of the content of messages or information delivered via the intervention, are crucial. Expertise was evident in most of the interventions (n=10, 33%), as the content was created by experts. Multidisciplinary teams consisting of researchers, physicians, clinicians, therapists, and practitioners (ie, software developers) were involved in creating content for the interventions [17,19-21,23-25,31,36,43]. This was followed by authority (n=9, 30%) and verification (n=50, 17%). Authority was identified in several interventions [17,19-21,25,30,33,36,43]. This was in the form of citing health quotes from recognized authorities such as the National Heart Foundation of Australia and the American College for Sports Medicine. Verifiability was identified in 17% (n=5) of the interventions [22,23,25,27,33]. In these interventions, links were provided for fact-checking

purposes. A total of 5 (17%) interventions were observed to implement the "real-world feel" feature such that participants could contact researchers or clinicians for various purposes [22-24,31,42].

Social Support

Social learning was the only social support feature identified in the analyzed interventions. The intervention enabled users to watch others go through a similar behavior change process [38]. Other features, such as social comparison, cooperation, social facilitation, normative influence, competition, and recognition, were not identified in the research articles. The minimal use of social support features may be associated with the tendency of these features to trigger high intensities of negative sentiment and emotional backfire [48,49].

Discussion

Principal Findings and Implications

CHD is a severe health problem, with an increasing prevalence worldwide. Adopting a healthy lifestyle and being aware of CHD risk factors are essential for its prevention and management. This scoping review identified existing health BCSSs from RCTs, reported their characteristics, and analyzed the persuasion context and persuasive design principles of the systems identified for CHD self-management using the PSD model. We found that trends in the use of persuasive system features on par with behavior change theories were identified for only 50% (n=15) of the RCTs; this points to the fact that there is still no consensus on the need to use theories and the best approaches to design interventions to promote behavior change in patients with CHD.

The intention behind the BCSSs was to support patients with CHD to improve their health and lifestyle. The interventions analyzed sought to improve clinical outcomes, behavioral outcomes, psychological outcomes, and improved quality of life outcomes associated with CHD for their users. The analysis of the events provided insights into the problem domain, user characteristics, and technology-related factors. Although the main user group was people who had been diagnosed with CHD, some of them had multiple health conditions such as diabetes and depression or needed to change at least one behavior (eg, smoking, physical inactivity, and medication adherence) to prevent further health complications. Users of the interventions were aged at least 18 years, cognitively sound, and computer literate.

Different types of information technology platforms were chosen for BCSS design and development for various reasons. For example, SMS and other types of text messages were used for delivering the intervention in 3% (n=1) of the studies because of their ease of use and cost-effectiveness. Moreover, WeChat, an instant messaging app, was adopted because it is a widely used social networking site in China.

Furthermore, the analysis of the strategy showed that the content delivered via the BCSSs was for educational purposes and hence required users to engage their cognitive resources. Another interesting finding was the lack of educational content in 13% (n=4) of the interventions. One would expect BCSSs to have educational content because such content provides a means to educate patients and reinforce behavior change [50].

Additionally, although the RCTs conducted in Europe used at least one persuasive feature in their interventions, none of the studies mentioned the behavior change theory that underpinned the BCSS. Proponents of the use of behavior change theories argue that theories explain the mechanisms through which behavior change occurs [51]. Often, developers prioritize the use of behavior change techniques at the expense of gaining an in-depth understanding of the theories that underlie them [14]. A potential problem that may arise with the use of BCSSs without a solid theoretical foundation is the creation of conflicting mechanisms that can affect the system's long-term effectiveness. This also points to the fact that there is still no consensus on the relevance of theories and best approaches to design interventions to promote behavior change in patients with CHD.

Another issue worth mentioning is designer or persuader bias and how it manifests itself in designing persuasive and behavior change interventions. This bias may have important ethical implications, as persuasive design decisions should be neither deceptive, manipulative, nor coercive [52]—nor cause harm to the users. Although a lot of emphasis has been placed on user-centered design in recent times, there may be a tendency for designers and developers to make design decisions affected by cognitive biases (such as the inability to evaluate all solutions to determine optimal solutions due to resource limitations) and illogical decisions (eg, intuitive reasoning, which uses low cognitive resources) [53]. In the design of BCSSs, persuader bias can manifest itself in, for example, the kind of content presented to users and the implementation of features. Designers are tasked to make design decisions based on insights generated from the context of use, user characteristics, and affordances of technology [13]. These decisions can influence the acceptability and, subsequently, the effectiveness of the system because users are sensitive to design features (eg, tailoring) [54]. From the results of the analysis, it appears that the involvement of users (eg, to create messages) may serve as a strategy to mitigate designer bias in the development of health interventions. Further research is needed to confirm this claim.

The findings from this review suggest that BCSSs have the potential to support the self-management of CHD and promote healthy lifestyles. Studies argue that the effectiveness of these systems depends on their design and implementation [55,56]. The major challenge that emerges is identifying the components of the BCSS and pertinent PSD features that are responsible for driving behavioral change [56]. This will require a deep understanding of the intricate interplay between various elements within the BCSS, such as behavior change theories, educational content, and persuasive features. Additionally, knowing the individual contribution of persuasive components and synergistic effects in this regard is desired [56]. Further research shows that the patterns of use of the intervention influence its effectiveness in achieving the desired outcome [57].

Apart from Putra et al [39], between 1 and 6 persuasive features were used in the analyzed BCSSs. In the primary task support category, personalization and self-monitoring were found to be

the most widely represented features. Within the dialogue support category, reminders and praise were the most-used features. Expertise and authority were the most-used features regarding credibility, while social learning was the only feature identified in the social support category. The frequent use of these features may imply their importance for BCSSs developed for managing or preventing CHD.

Additionally, we found that feedback is closely related to praise in the PSD model, yet distinct based on the textual descriptions. Users can receive feedback that is not necessarily of a positive tone (like praise). Feedback can be given on the user's performance, which may be motivational, constructive, or both. Another interesting finding was the use of the term "semipersonalization," a form of personalization. Semipersonalization describes a weak level of personalization. This brings to light the commentary by Oinas-Kukkonen [58] on a concept called "personalization myopia," which seeks to clarify the misunderstanding surrounding the level and type of personalization offered in mobile and web apps. The level and type of personalization identified in the studies varied. In the studies analyzed, personalization of content was widely used. For instance, mobile text message–based interventions simply used the names of users to indicate personalization. With that being said, some of the studies did clarify the use of the level of personalization by using semipersonalization. A critical look at some studies revealed the use of user data to generate personalized content other than just the name. This is a step in the right direction toward true personalization. True personalization requires detailed information, such as the user's preferences, to create individualized experiences [59] in the different stages of the user's journey through the app. Moreover, the use of personalization accounts for differences in preferences between and within groups of participants [60].

It was reassuring that design principles that make the system credible were identified in 60% (n=18) of the BCSSs studied. System credibility support principles tend to increase user satisfaction and influence their intention to use an intervention [61]. Incorporating this principle shows the intention and commitment of the designers to deliver credible content and a trustworthy system.

Limitations

We acknowledge some limitations in our work. First, we relied on textual descriptions provided in the original papers, which were subjectively interpreted by the reviewers, creating an avenue for subjective bias. Second, our study is based on English-language publications only, thus excluding publications in other languages that could provide rich content to this analysis. Third, this review did not analyze the effectiveness of the analyzed interventions as it fell beyond the scope of this research. Instead, the focus was on identifying trends in persuasive design characteristics used in BCSSs for CHD. This encompassed key persuasive elements like behavior change theories, techniques, educational content, and persuasive features in the analyzed studies. Also, recognizing prevailing trends can contribute to the refinement and development of theories and practices for BCSSs. This valuable information can also help developers, clinicians, and developers to make informed decisions. To deepen our understanding, future research should assess the relevance and impact of persuasive elements through meta-analysis. This approach may yield much-needed evidence and insights and contribute to developing effective interventions.

Conclusion

This study sought to identify persuasive design characteristics in mobile, web, and other information system interventions for CHD. We analyzed 30 peer-reviewed RCT papers that implemented BCSSs for patients with CHD. This study highlights the key issues that should be considered when designing and developing BCSSs for CHD. From the analysis of RCTs, we found that trends in the use of persuasive system features on par with behavior change theories were identified for only 50% (n=15) of RCTs. This points to the fact that there is still no consensus on the need to use theories and the best approaches to design interventions to promote behavior change in patients with CHD. Although we were able to highlight the trends in the persuasive features, we were unable to determine if these features influenced the effectiveness of the analyzed BCSSs. Moreover, there is a need to evaluate the actual effect of the intervention on users. Thus, future research should address this using data analysis methodologies such as meta-analysis.

Acknowledgments

We would like to thank Markku Savolainen, MD and Piiastiina Tikka, PhD, for their feedback and comments on an early draft of this paper. Additionally, we would like to express our gratitude to Sami Pohjolainen for his thorough review of the Persuasive Systems Design features. This study has received funding from the European Union's Horizon 2020 research and innovation program (grant 848056) and from the Academy of Finland project, known as PerFeat (decision 351670). It reflects only the authors' views and neither the Commission of Finland nor the Academy of Finland are responsible for any use that may be made of the information it contains.

Data Availability

The data used in this research is publicly available in the Multimedia Appendices.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The analysis of intent describes the primary and secondary outcomes of the randomized controlled trials.
[XLSX File (Microsoft Excel File), 18 KB - cardio_v8i1e49515_app1.xlsx]

Multimedia Appendix 2
The analysis of the event describes the use, user, and technology context in the studied interventions.
[XLSX File (Microsoft Excel File), 18 KB - cardio_v8i1e49515_app2.xlsx]

Multimedia Appendix 3
The analysis of the strategy describes the message and route for presenting information in the studied intervention.
[XLSX File (Microsoft Excel File), 17 KB - cardio_v8i1e49515_app3.xlsx]

Multimedia Appendix 4
Distribution of identified persuasive systems features, contents, and theories as shown by the shaded areas. "Content" refers to the number of contents identified in each intervention. "Theories" refers to the number of theories identified in each intervention. AUT: authority; CRED: credibility support; DIAL: dialogue support; EXP: expertise; PER: personalization; PRA: praise; PRIM: primary task support; RED: reduction; REM: reminder; REW: rewards; RWF: real-world feel; SMO: self-monitoring; SOCI, SLE: social learning; SRO: social role; SUG: suggestion; TAI: tailoring; VER: verifiability.
[XLSX File (Microsoft Excel File), 12 KB - cardio_v8i1e49515_app4.xlsx]

Multimedia Appendix 5
The analysis of persuasive features used in the interventions examined.
[XLSX File (Microsoft Excel File), 21 KB - cardio_v8i1e49515_app5.xlsx]

Multimedia Appendix 6
PRISMA checklist.
[PDF File (Adobe PDF File), 120 KB - cardio_v8i1e49515_app6.pdf]

References

1. Tsao CW, Aday AW, Almarzooq ZI, Anderson CA, Arora P, Avery CL, et al. Heart disease and stroke statistics—2023 update: A report From the American Heart Association. Circulation 2023 Feb 21;147(8). [doi: 10.1161/cir.0000000000001123]
2. Jung HG, Yang YK. Factors influencing health behavior practice in patients with coronary artery diseases. Health Qual Life Outcomes 2021 Jan 06;19(1):3 [FREE Full text] [doi: 10.1186/s12955-020-01635-2] [Medline: 33407579]
3. Suls J, Mogavero JN, Falzon L, Pescatello LS, Hennessy EA, Davidson KW. Health behaviour change in cardiovascular disease prevention and management: meta-review of behaviour change techniques to affect self-regulation. Health Psychol Rev 2020 Mar 29;14(1):43-65 [FREE Full text] [doi: 10.1080/17437199.2019.1691622] [Medline: 31707938]
4. Oinas-Kukkonen H. Behavior change support systems: A research model and agenda. In: Ploug T, Hasle P, Oinas-Kukkonen H, editors. Persuasive Technology. PERSUASIVE 2010. Lecture Notes in Computer Science, vol 6137. Berlin, Germany: Springer; 2010:4-14.
5. Oinas-Kukkonen H. A foundation for the study of behavior change support systems. Pers Ubiquit Comput 2012 Jul 19;17(6):1223-1235. [doi: 10.1007/s00779-012-0591-5]
6. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
7. Cruz-Cobo C, Bernal-Jiménez MÁ, 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 Dec 02;10(12):e39593 [FREE Full text] [doi: 10.2196/39593] [Medline: 36459396]
8. Indraratna P, Tardo D, Yu J, Delbaere K, Brodie M, Lovell N, et al. Mobile phone technologies in the management of ischemic heart disease, heart failure, and hypertension: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e16695 [FREE Full text] [doi: 10.2196/16695] [Medline: 32628615]
9. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013 Feb 20;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
10. Subedi N, Rawstorn JC, Gao L, Koorts H, Maddison R. Implementation of telerehabilitation interventions for the self-management of cardiovascular disease: systematic review. JMIR Mhealth Uhealth 2020 Nov 27;8(11):e17957 [FREE Full text] [doi: 10.2196/17957] [Medline: 33245286]

11. Bohlen L, Michie S, de Bruin M, Rothman A, Kelly M, Groarke H, et al. Do combinations of behavior change techniques that occur frequently in interventions reflect underlying theory? Ann Behav Med 2020 Nov 01;54(11):827-842 [FREE Full text] [doi: 10.1093/abm/kaaa078] [Medline: 32959875]
12. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020 Mar 18;8(3):e17046 [FREE Full text] [doi: 10.2196/17046] [Medline: 32186518]
13. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Sys 2009;24. [doi: 10.17705/1CAIS.02428]
14. Spohrer K, Fallon M, Hoehle H, Heinzl A. Designing effective mobile health apps: does combining behavior change techniques really create synergies? Journal of Management Information Systems 2021 Aug 06;38(2):517-545. [doi: 10.1080/07421222.2021.1912936]
15. Merz M, Ackermann L. Design principles of persuasive systems – review and discussion of the persuasive systems design model. In: AMCIS 2021 Proceedings. 2021 Presented at: Americas Conference on Information Systems (AMCIS); August 9-13, 2021; Virtual conference p. 3 URL: https://aisel.aisnet.org/amcis2021/sig_hci/sig_hci/3
16. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
17. Shariful Islam SM, Chow CK, Redfern J, Kok C, Rådholm K, Stepien S, et al. Effect of text messaging on depression in patients with coronary heart disease: a substudy analysis from the TEXT ME randomised controlled trial. BMJ Open 2019 Feb 20;9(2):e022637 [FREE Full text] [doi: 10.1136/bmjopen-2018-022637] [Medline: 30787075]
18. Devi R, Powell J, Singh S. A web-based program improves physical activity outcomes in a primary care angina population: randomized controlled trial. J Med Internet Res 2014 Sep 12;16(9):e186 [FREE Full text] [doi: 10.2196/jmir.3340] [Medline: 25217464]
19. Huo X, Krumholz HM, Bai X, Spatz ES, Ding Q, Horak P, et al. Effects of mobile text messaging on glycemic control in patients with coronary heart disease and diabetes mellitus: a randomized clinical trial. Circ Cardiovasc Qual Outcomes 2019 Sep;12(9):e005805. [doi: 10.1161/CIRCOUTCOMES.119.005805] [Medline: 31474119]
20. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. JAMA 2015 Sep;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
21. Zheng X, Spatz ES, Bai X, Huo X, Ding Q, Horak P, et al. Effect of text messaging on risk factor management in patients with coronary heart disease: the CHAT randomized clinical trial. Circ Cardiovasc Qual Outcomes 2019 Apr;12(4):e005616. [doi: 10.1161/CIRCOUTCOMES.119.005616] [Medline: 30998400]
22. Vernooij JWP, Kaasjager HAH, van der Graaf Y, Wierdsma J, Grandjean HMH, Hovens MMC, SMARTStudy Group. Internet based vascular risk factor management for patients with clinically manifest vascular disease: randomised controlled trial. BMJ 2012 Jun 12;344:e3750 [FREE Full text] [doi: 10.1136/bmj.e3750] [Medline: 22692651]
23. Zhang H, Jiang Y, Nguyen HD, Poo DCC, Wang W. The effect of a smartphone-based coronary heart disease prevention (SBCHDP) programme on awareness and knowledge of CHD, stress, and cardiac-related lifestyle behaviours among the working population in Singapore: a pilot randomised controlled trial. Health Qual Life Outcomes 2017 Mar 14;15(1):49 [FREE Full text] [doi: 10.1186/s12955-017-0623-y] [Medline: 28288636]
24. Reid RD, Morrin LI, Beaton LJ, Papadakis S, Kocourek J, McDonnell L, et al. Randomized trial of an internet-based computer-tailored expert system for physical activity in patients with heart disease. Eur J Prev Cardiol 2012 Dec;19(6):1357-1364. [doi: 10.1177/1741826711422988] [Medline: 21903744]
25. Bae J, Woo S, Lee J, Park S, Kwon SW, Choi SH, et al. mHealth interventions for lifestyle and risk factor modification in coronary heart disease: randomized controlled trial. JMIR Mhealth Uhealth 2021 Sep 24;9(9):e29928 [FREE Full text] [doi: 10.2196/29928] [Medline: 34559058]
26. Pfaeffli Dale L, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R. Text message and internet support for coronary heart disease self-management: results from the Text4Heart randomized controlled trial. J Med Internet Res 2015 Oct 21;17(10):e237 [FREE Full text] [doi: 10.2196/jmir.4944] [Medline: 26490012]
27. Dorje T, Zhao G, Tso K, Wang J, Chen Y, Tsokey L, et al. Smartphone and social media-based cardiac rehabilitation and secondary prevention in China (SMART-CR/SP): a parallel-group, single-blind, randomised controlled trial. Lancet Digit 2019 Nov;1(7):e363-e374. [doi: 10.1016/s2589-7500(19)30151-7]
28. Duan YP, Liang W, Guo L, Wienert J, Si GY, Lippke S. Evaluation of a web-based intervention for multiple health behavior changes in patients with coronary heart disease in home-based rehabilitation: pilot randomized controlled trial. J Med Internet Res 2018 Nov 19;20(11):e12052 [FREE Full text] [doi: 10.2196/12052] [Medline: 30455167]
29. Frederix I, Van Driessche N, Hansen D, Berger J, Bonne K, Alders T, et al. Increasing the medium-term clinical benefits of hospital-based cardiac rehabilitation by physical activity telemonitoring in coronary artery disease patients. Eur J Prev Cardiol 2015 Feb;22(2):150-158. [doi: 10.1177/2047487313514018] [Medline: 24249840]

30. Indraratna P, Biswas U, McVeigh J, Mamo A, Magdy J, Vickers D, et al. A smartphone-based model of care to support patients with cardiac disease transitioning from hospital to the community (TeleClinical Care): pilot randomized controlled trial. JMIR Mhealth Uhealth 2022 Feb 28;10(2):e32554 [FREE Full text] [doi: 10.2196/32554] [Medline: 35225819]
31. Kamal AK, Khalid W, Muqeet A, Jamil A, Farhat K, Gillani SRA, et al. Making prescriptions "talk" to stroke and heart attack survivors to improve adherence: Results of a randomized clinical trial (The Talking Rx Study). PLoS One 2018;13(12):e0197671 [FREE Full text] [doi: 10.1371/journal.pone.0197671] [Medline: 30571697]
32. 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 Feb 15;10(2):e32251 [FREE Full text] [doi: 10.2196/32251] [Medline: 34906924]
33. Maddison R, Pfaeffli L, Whittaker R, Stewart R, Kerr A, Jiang Y, et al. A mobile phone intervention increases physical activity in people with cardiovascular disease: Results from the HEART randomized controlled trial. Eur J Prev Cardiol 2015 Jun;22(6):701-709. [doi: 10.1177/2047487314535076] [Medline: 24817694]
34. Maddison R, Rawstorn JC, Stewart RAH, Benatar J, Whittaker R, Rolleston A, et al. Effects and costs of real-time cardiac telerehabilitation: randomised controlled non-inferiority trial. Heart 2019 Jan;105(2):122-129 [FREE Full text] [doi: 10.1136/heartjnl-2018-313189] [Medline: 30150328]
35. Maddison R, Jiang Y, Stewart R, Scott T, Kerr A, Whittaker R, et al. An intervention to improve medication adherence in people with heart disease (Text4HeartII): randomized controlled trial. JMIR Mhealth Uhealth 2021 Jun 09;9(6):e24952 [FREE Full text] [doi: 10.2196/24952] [Medline: 34106081]
36. Ni Z, Wu B, Yang Q, Yan LL, Liu C, Shaw RJ. An mHealth intervention to improve medication adherence and health outcomes among patients with coronary heart disease: randomized controlled trial. J Med Internet Res 2022 Mar 09;24(3):e27202 [FREE Full text] [doi: 10.2196/27202] [Medline: 35262490]
37. Park LG, Howie-Esquivel J, Chung ML, Dracup K. A text messaging intervention to promote medication adherence for patients with coronary heart disease: a randomized controlled trial. Patient Educ Couns 2014 Feb;94(2):261-268. [doi: 10.1016/j.pec.2013.10.027] [Medline: 24321403]
38. Pfaeffli Dale L, Whittaker R, Dixon R, Stewart R, Jiang Y, Carter K, et al. Acceptability of a mobile health exercise-based cardiac rehabilitation intervention: a randomized trial. J Cardiopulm Rehabil Prev 2015;35(5):312-319. [doi: 10.1097/HCR.0000000000000125] [Medline: 26181037]
39. Putra R, Rochmawati I, Prima D. The impact of mobile application: "Friends of Heart" in knowledge and compliance of patients with coronary heart disease. Pharm Educ 2022;22(2):36-40. [doi: 10.46542/pe.2022.222.3640]
40. Silarova B, Sharp S, Usher-Smith JA, Lucas J, Payne RA, Shefer G, et al. Effect of communicating phenotypic and genetic risk of coronary heart disease alongside web-based lifestyle advice: the INFORM Randomised Controlled Trial. Heart 2019 Jul 30;105(13):982-989 [FREE Full text] [doi: 10.1136/heartjnl-2018-314211] [Medline: 30928969]
41. Snoek JA, Prescott EI, van der Velde AE, Eijsvogels TMH, Mikkelsen N, Prins LF, et al. Effectiveness of home-based mobile guided cardiac rehabilitation as alternative strategy for nonparticipation in clinic-based cardiac rehabilitation among elderly patients in Europe: a randomized clinical trial. JAMA Cardiol 2021 Apr 01;6(4):463-468 [FREE Full text] [doi: 10.1001/jamacardio.2020.5218] [Medline: 33112363]
42. Song Y, Ren C, Liu P, Tao L, Zhao W, Gao W. Effect of smartphone-based telemonitored exercise rehabilitation among patients with coronary heart disease. J Cardiovasc Transl Res 2020 Aug;13(4):659-667 [FREE Full text] [doi: 10.1007/s12265-019-09938-6] [Medline: 31820334]
43. Thakkar J, Redfern J, Thiagalingam A, Chow CK. Patterns, predictors and effects of texting intervention on physical activity in CHD - insights from the TEXT ME randomized clinical trial. Eur J Prev Cardiol 2016 Nov 20;23(17):1894-1902. [doi: 10.1177/2047487316664190] [Medline: 27512051]
44. Torri A, Panzarino C, Scaglione A, Modica M, Bordoni B, Redaelli R, et al. Promotion of home-based exercise training as secondary prevention of coronary heart disease: a pilot web-based intervention. J Cardiopulm Rehabil Prev 2018 Jul;38(4):253-258. [doi: 10.1097/HCR.0000000000000316] [Medline: 29738378]
45. Yudi MB, Clark DJ, Tsang D, Jelinek M, Kalten K, Joshi S, et al. SMARTphone-based, early cardiac REHABilitation in patients with acute coronary syndromes: a randomized controlled trial. Coron Artery Dis 2021 Aug 01;32(5):432-440. [doi: 10.1097/MCA.0000000000000938] [Medline: 32868661]
46. Santo K, Singleton A, Rogers K, Thiagalingam A, Chalmers J, Chow CK, et al. Medication reminder applications to improve adherence in coronary heart disease: a randomised clinical trial. Heart 2019 Feb 27;105(4):323-329. [doi: 10.1136/heartjnl-2018-313479] [Medline: 30150326]
47. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: a systematic review. J Med Syst 2016 Mar;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
48. Karppinen P, Oinas-Kukkonen H, Alahäivälä T, Jokelainen T, Keränen AM, Salonurmi T, et al. Persuasive user experiences of a health behavior change support system: A 12-month study for prevention of metabolic syndrome. Int J Med Inform 2016 Dec;96:51-61. [doi: 10.1016/j.ijmedinf.2016.02.005] [Medline: 26992482]
49. Ekpezu AO, Wiafe I, Oinas-Kukkonen H. Enhancing Perceived Health Competence: The Impact of Persuasive Social Support Features in Health and Fitness Apps. Int J Hum–Comp Interact 2023 Nov 21:1-15. [doi: 10.1080/10447318.2023.2277493]

50. Bashi N, Fatehi F, Fallah M, Walters D, Karunanithi M. Self-management education through mhealth: review of strategies and structures. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e10771 [FREE Full text] [doi: 10.2196/10771] [Medline: 30341042]
51. Connell LE, Carey RN, de Bruin M, Rothman AJ, Johnston M, Kelly MP, et al. Links between behavior change techniques and mechanisms of action: an expert consensus study. Ann Behav Med 2019 Jul 17;53(8):708-720 [FREE Full text] [doi: 10.1093/abm/kay082] [Medline: 30452535]
52. Smids J. The voluntariness of persuasive technology. In: Bang M, Ragnemalm EL, editors. Persuasive Technology. Design for Health and Safety. PERSUASIVE 2012. Lecture Notes in Computer Science, vol 7284. Berlin, Germany: Springer; 2012:123-132.
53. Tang A. Software designers, are you biased? In: SHARK '11: Proceedings of the 6th International Workshop on SHAring and Reusing Architectural Knowledge. New York, NY: Association for Computing Machinery; 2011:1-8.
54. Materia FT, Smyth JM. Acceptability of intervention design factors in mHealth intervention research: experimental factorial study. JMIR Mhealth Uhealth 2021 Jul 26;9(7):e23303 [FREE Full text] [doi: 10.2196/23303] [Medline: 34309563]
55. Agyei E, Miettunen J, Oinas-Kukkonen H. Effective interventions and features for coronary heart disease: a meta-analysis. Behaviour & Information Technology 2023 May 19:1-17. [doi: 10.1080/0144929x.2023.2213342]
56. Goodwin L, Ostuzzi G, Khan N, Hotopf MH, Moss-Morris R. Can we identify the active ingredients of behaviour change interventions for coronary heart disease patients? a systematic review and meta-analysis. PLoS One 2016;11(4):e0153271 [FREE Full text] [doi: 10.1371/journal.pone.0153271] [Medline: 27105435]
57. Agyei E, Oinas-Kukkonen H, Nyman V, Virkkula T, Oikarinen N, Merikallio H, et al. Assessing interventional components in a weight loss app. 2023 Presented at: 56th Hawaii International Conference on System Sciences; January 3-6, 2023; Maui, HI p. 2953-2962 URL: https://hdl.handle.net/10125/102994 [doi: 10.24251/hicss.2023.363]
58. Oinas-Kukkonen H. Personalization myopia: a viewpoint to true personalization of information systems. In: Mindtrek '18: Proceedings of the 22nd International Academic Mindtrek Conference. New York, NY: Association for Computing Machinery; 2018:88-91.
59. Oinas-Kukkonen H, Pohjolainen S, Agyei E. Mitigating issues with/of/for true personalization. Front Artif Intell 2022 Apr 26;5:844817 [FREE Full text] [doi: 10.3389/frai.2022.844817] [Medline: 35558170]
60. Chatterjee A, Prinz A, Gerdes M, Martinez S. Digital interventions on healthy lifestyle management: systematic review. J Med Internet Res 2021 Nov 17;23(11):e26931 [FREE Full text] [doi: 10.2196/26931] [Medline: 34787575]
61. Handayani PW, Gelshirani NB, Azzahro F, Pinem AA, Hidayanto AN. The influence of argument quality, source credibility, and health consciousness on satisfaction, use intention, and loyalty on mobile health application use. Informatics in Medicine Unlocked 2020;20:100429. [doi: 10.1016/j.imu.2020.100429]

Abbreviations

BCSS: behavior change support system
CAD: coronary artery disease
CHD: coronary heart disease
PICO: Population, Intervention, Comparison, and Outcome
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PSD: Persuasive Systems Design
RCT: randomized controlled trial

Edited by A Mavragani; submitted 02.06.23; peer-reviewed by W Liang, Q Ding; comments to author 19.07.23; revised version received 21.08.23; accepted 01.05.24; published 19.06.24.

Please cite as:
Agyei EEYF, Ekpezu A, Oinas-Kukkonen H
Persuasive Systems Design Trends in Coronary Heart Disease Management: Scoping Review of Randomized Controlled Trials
JMIR Cardio 2024;8:e49515
URL: https://cardio.jmir.org/2024/1/e49515
doi:10.2196/49515
PMID:

Original Paper

Feasibility of Using Text Messaging to Identify and Assist Patients With Hypertension With Health-Related Social Needs: Cross-Sectional Study

Aryn Kormanis[1], DO; Selina Quinones[1], BS; Corey Obermiller[1], MAS; Nancy Denizard-Thompson[1], MD; Deepak Palakshappa[1,2,3], MS, MD

[1]Department of Internal Medicine, Wake Forest University School of Medicine, Winston Salem, NC, United States

[2]Department of Pediatrics, Wake Forest University School of Medicine, Winston Salem, NC, United States

[3]Department of Epidemiology and Prevention, Wake Forest University School of Medicine, Winston Salem, NC, United States

Corresponding Author:
Deepak Palakshappa, MS, MD
Department of Internal Medicine
Wake Forest University School of Medicine
Medical Center Blvd
Winston Salem, NC, 27157
United States
Phone: 1 3367161795
Email: dpalaksh@wakehealth.edu

Abstract

Background: Health-related social needs are associated with poor health outcomes, increased acute health care use, and impaired chronic disease management. Given these negative outcomes, an increasing number of national health care organizations have recommended that the health system screen and address unmet health-related social needs as a routine part of clinical care, but there are limited data on how to implement social needs screening in clinical settings to improve the management of chronic diseases such as hypertension. SMS text messaging could be an effective and efficient approach to screen patients; however, there are limited data on the feasibility of using it.

Objective: We conducted a cross-sectional study of patients with hypertension to determine the feasibility of using SMS text messaging to screen patients for unmet health-related social needs.

Methods: We randomly selected 200 patients (≥18 years) from 1 academic health system. Patients were included if they were seen at one of 17 primary care clinics that were part of the academic health system and located in Forsyth County, North Carolina. We limited the sample to patients seen in one of these clinics to provide tailored information about local community-based resources. To ensure that the participants were still patients within the clinic, we only included those who had a visit in the previous 3 months. The SMS text message included a link to 6 questions regarding food, housing, and transportation. Patients who screened positive and were interested received a subsequent message with information about local resources. We assessed the proportion of patients who completed the questions. We also evaluated for the differences in the demographics between patients who completed the questions and those who did not using bivariate analyses.

Results: Of the 200 patients, the majority were female (n=109, 54.5%), non-Hispanic White (n=114, 57.0%), and received commercial insurance (n=105, 52.5%). There were no significant differences in demographics between the 4446 patients who were eligible and the 200 randomly selected patients. Of the 200 patients included, the SMS text message was unable to be delivered to 9 (4.5%) patients and 17 (8.5%) completed the social needs questionnaire. We did not observe a significant difference in the demographic characteristics of patients who did versus did not complete the questionnaire. Of the 17, a total of 5 (29.4%) reported at least 1 unmet need, but only 2 chose to receive resource information.

Conclusions: We found that only 8.5% (n=17) of patients completed a SMS text message–based health-related social needs questionnaire. SMS text messaging may not be feasible as a single modality to screen patients in this population. Future research should evaluate if SMS text message–based social needs screening is feasible in other populations or effective when paired with other screening modalities.

(JMIR Cardio 2024;8:e54530) doi:10.2196/54530

KEYWORDS

social determinants of health; health-related social needs; mobile health; health information technology; feasibility; mobile phone; SMS text messaging; message; pilot study; patients; patient; hypertension; screening

Introduction

Unmet health-related social needs, such as food insecurity and housing instability, are associated with impaired chronic disease management and worse health outcomes [1-5]. For example, people with hypertension who live in a food-insecure household are more likely to have worse diet quality and blood pressure control than people with hypertension who live in a food-secure household. Because of their negative impact, national organizations, such as the Centers for Medicare and Medicaid (CMS), have recommended that health systems integrate interventions to screen and address patients' unmet social needs as a routine part of clinical care [6-9]. Although there has been growing investment by health systems to integrate these interventions, there are still limited data on how to most effectively implement screening in busy clinical settings [10,11]. Studies assessing the use of mobile tools (eg, tablets) and telephone-based screening identified barriers to these screening modalities [12-14]. Barriers to using these methods include that tablets are dependent on patients presenting in person to the clinic [12,15,16] and phone-based screening may add additional work for already busy clinical staff [14].

SMS text messaging could be an effective and efficient approach to assess patients for health-related social needs and allow for screening to occur outside of the direct patient encounter. However, prior studies have not used SMS text messaging to screen patients for health-related social needs. To fill this gap, our objective was to determine the feasibility and acceptability of using SMS text messaging to screen and assist patients with hypertension with health-related social needs. We were specifically interested in evaluating if patients would complete a SMS text message–based social risk questionnaire and if there were differences in demographics between patients who completed the questionnaire and those who did not.

Methods

Study Design and Population

We conducted a pilot cross-sectional study at Atrium Health Wake Forest Baptist Health (AHWFB) to assess the feasibility of using SMS text messaging to send a link to a web-based questionnaire to screen patients for health-related social needs. AHWFB is a large, integrated academic health system serving communities in Central and Western North Carolina. The system is comprised of a tertiary care hospital located in Winston-Salem, NC, 4 community hospitals, and >300 ambulatory practices that all use a single electronic health record (EHR; EpicCare). We identified eligible adult patients (≥18 years) with hypertension who were seen at an AHWFB primary care clinic in the previous 3 months (between November 2022 and February 2023). We only included patients who had been seen in the last 3 months to ensure they were still a patient at the clinic. We limited the sample to patients seen in an AHWFB internal or family medicine clinic (17 clinics) in Forsyth County, North Carolina to provide tailored information about local community-based resources (eg, food pantries). We also included these 17 clinics because they are in the process of integrating social risk screening into routine clinical care, but none of the 17 clinics had implemented a standardized screening process prior to or during the time period the study was conducted. Based on 2023 census estimates of people living in Forsyth County, 17.2% of the population are older than 65 years of age, 65.7% identify as non-Hispanic White, 27.7% as non-Hispanic Black, and 14.3% as Hispanic. A total of 14% of the population have a household income below the federal poverty level, 12.9% of the population are estimated to be uninsured, and 11.6% of households in the county are estimated to be food insecure. We recruited participants by taking a simple random sample of 200 patients from the 4446 eligible patients identified and sending them SMS text messages to the cell phone number included in the EHR. As there is no specific consensus on the sample size necessary to assess the feasibility of a pilot study, consistent with prior studies, we included 200 (~5% of the eligible population) patients [17,18].

Ethical Considerations

The Wake Forest University School of Medicine institutional review board reviewed and approved this study under the expedited review with a waiver of written informed consent (IRB00092658). The SMS text message included language to notify the participants that the questionnaire was part of a research study and that participation was voluntary. To maintain privacy and confidentiality, all participants' responses to the social risk questionnaire and personal health and demographic data were stored in a password-protected file on the institution's secure server. Only study team members had access to the data. The participants did not receive compensation for participation in the study.

SMS Text Messages

We developed the SMS text message–based on a detailed literature review of the social needs literature (Multimedia Appendix 1). The AHWFB digital communication committee also reviewed the SMS text messages. The committee includes patients, clinicians, hospital administrators, and members of the institutional review board. The committee reviews all SMS text messages that may be sent to patients for either research or clinical purposes at the institution, and they provide input on the wording of the messages. We also reviewed the SMS text messages with patients who were not included in the study to assess face validity and to provide additional input on the messages. The SMS text message included a link to a questionnaire with 6 questions from the CMS Accountable Health Communities Health-Related Social Needs Screening Tool [7]. The 6 questions included 2 questions about housing (including 1 about patients' current housing situation and 1 about problems with housing), 2 questions about food insecurity, 1 question about transportation, and 1 question about use. We limited to questions regarding food, housing, transportation,

and utilities because there were local resources available to assist patients with these domains. The initial message was sent to the patient's cell phone number listed in the EHR on March 14, 2023. All of the 4446 patients eligible had a phone number listed in the EHR. For those that did not respond, the same message was sent again 1 week later. Responses were documented in the EHR. The SMS text message and questionnaire were sent in English or Spanish based on the preferred language of the patient listed in the EHR. We used the standard scoring to identify patients who screened positive for a social need. The patients were identified as having housing instability if they provided a response other than "I have a steady place to live" to the first housing question or if they provided a response other than "None of the above" to the second question. Patients were identified as having food insecurity if they responded to either of the 2 questions with "sometimes true" or "often true." Patients were identified as having a lack of transportation if they responded "yes" to the transportation question, and they were identified as having difficulty with utilities if they responded "yes" or "already shut off" to the utilities question.

For patients who screened positive for any of the social needs, they were asked if they would be interested in receiving information about local community resources. A list of resources was then sent by the system in a subsequent SMS text message to patients who were interested in receiving information about local community resources. The system used branching logic and only provided information on resources for the unmet need the patient reported (eg, food resources for those with food insecurity). Patients were asked if they would be willing to receive an additional message 1 month later to assess the acceptability of the process and if they used any of the information provided about community resources (Multimedia Appendix 1). Questions were based on the validated Acceptability of Intervention Measure [19].

Statistical Analysis

We obtained age, sex, race (White, Black, American Indian or Alaska Native, Asian Indian, Filipino, and other), ethnicity (Hispanic and non-Hispanic), insurance status (commercial or private, Medicaid, Medicare, uninsured, and other), and preferred language (English or Spanish) for all patients through data extraction from the EHR. Informed by the Reach, Effectiveness, Acceptability, Implementation, and Maintenance (RE-AIM) framework [20-22], feasibility was based on the reach of the screening or the proportion of patients who completed the social needs questionnaire. To understand if using a SMS text message–based social needs questionnaire could lead to disparities in who completes screening questions, we evaluated for differences in demographics between patients who completed the questions and those who did not use bivariate analyses. We used either the chi-square or Fisher exact test for categorical variables, and we used the Welch *t* test for continuous variables to account for unequal sample sizes and variances. We considered an α of <.05 significant and all analyses were conducted using Stata 15.0 (StataCorp).

Results

Of the 200 patients randomly selected, the majority were females (n=109, 54.5%), non-Hispanic White (n=114, 57.0%), who received commercial or private health insurance (n=105, 52.5%), and had English listed as their preferred language (n=192, 96.0%). The mean age of the patients selected was 57.6 (SD 12.9) years of age.

Of the 200 patients, the SMS text message was unable to be delivered to 9 patients (either because the number listed was no longer working or was not a cell phone). A total of 17 (8.5%) patients completed the social needs questionnaire (Table 1).

We did not find a significant difference in demographics between patients who did and those who did not complete the questionnaire. Of the 17 who completed the questionnaire, 5 (29.4%) reported at least 1 unmet social need, but only 2 chose to receive local resource information. One person completed the follow-up questionnaire and reported that they learned about new community resources through the process.

Table 1. Study patient characteristics[a].

Characteristics	Total (N=200)	Did not respond (n=183)	Responded (n=17)	*P* value
Age (years), mean (SD)	57.6 (12.9)	57.1 (13.0)	63.1 (11.6)	.05
Sex, n (%)				.16
Male	91 (45.5)	86 (47.0)	5 (29.4)	
Female	109 (54.5)	97 (53.0)	12 (70.6)	
Race, n (%)				.99
Non-Hispanic and White	114 (57.0)	104 (56.8)	10 (58.8)	
Non-Hispanic and Black	67 (33.5)	61 (33.3)	6 (35.3)	
American Indian or Alaska Native	1 (0.5)	1 (0.6)	0 (0.0)	
Asian Indian	3 (1.5)	3 (1.6)	0 (0.0)	
Filipino	1 (0.5)	1 (0.6)	0 (0.0)	
Other	14 (7.0)	13 (7.1)	1 (5.9)	
Ethnicity, n (%)				.99
Hispanic, Latino, or Spanish	15 (7.5)	14 (7.7)	1 (5.9)	
Not Hispanic, Latino, or Spanish	185 (92.5)	169 (92.4)	16 (94.1)	
Health insurance, n (%)				.17
Commercial	105 (52.5)	100 (54.6)	5 (29.4)	
Medicaid	13 (6.5)	12 (6.6)	1 (5.9)	
Medicare	64 (32.0)	55 (30.1)	9 (52.9)	
Uninsured	11 (5.5)	10 (5.5)	1 (5.9)	
Other	7 (3.5)	6 (3.3)	1 (5.9)	
Language, n (%)				.52
English	192 (96.0)	176 (96.2)	16 (94.1)	
Spanish	8 (4.0)	7 (3.8)	1 (5.9)	
Food insecurity, n (%)				
Yes	N/A[b]	N/A	3 (17.7)	N/A
No	N/A	N/A	14 (82.4)	N/A
Living situation, n (%)				
I have a steady place to live	N/A	N/A	16 (94.1)	N/A
Prefer not answer	N/A	N/A	1 (5.9)	N/A
Problems in the home, n (%)				
None	N/A	N/A	16 (94.1)	N/A
Pests (eg, bugs, ants, or mice)	N/A	N/A	1 (5.9)	N/A
Lack of transportation, n (%)				
No	N/A	N/A	14 (87.5)	N/A
Yes	N/A	N/A	1 (6.3)	N/A
Prefer not answer	N/A	N/A	1 (6.3)	N/A
Electric, gas, or water shut off, n (%)				
No	N/A	N/A	17 (100.0)	N/A

[a]Bivariate analysis comparing characteristics of patients who were did and did not respond to a SMS text message linked social needs questionnaire in March 2023; responses to the social needs questionnaire for the 17 participants are also included.

[b]N/A: not applicable.

Discussion

These results suggest that SMS text messaging may be inadequate when used as a single modality for screening patients for unmet health-related social needs in this population, as only 17 (8.5%) of patients completed the social needs questionnaire. Yet 29% (n=5) of patients who completed the questionnaire reported having at least 1 unmet social need. Given the growing investment in integrating social care interventions into health care delivery, understanding screening strategies that are both effective and those that may be less effective are important.

There are several possible explanations for the low response rate. First, patients could have concerns about completing the social risk questionnaire using a SMS text message–based link. Prior studies have found that there are multiple factors, such as trust in their provider and concern about disclosure of sensitive information, contributing to patients' acceptability of social needs screening [23-25]. Previous studies have also found that patients are more likely to complete social risk questionnaires and disclose sensitive information if they are screened using paper or tablets in the clinic, rather than being verbally asked [26-29]. Patients may be unable to have confidence in who is administering the questionnaire and have access to the results using SMS text messaging. We did not notify patients that they would be sent the SMS text message. Discussing with patients prior to sending the message or directly tying the message to an upcoming visit may result in higher screening completion rates. It is also possible that the majority of patients who received the message did not have a social need, so they did not see a benefit in completing the social risk questionnaire.

A second possibility for the low response rate is the wording of the message. We tried to gather input from multiple different stakeholders in developing the message, but patients may have either misunderstood the purpose of the SMS text message or were not interested in participating in a research study. A SMS text message coming directly from a patient's provider or clinic could yield different results and future research could randomize who sends the message. A third reason for the low response rate could be barriers to using the technology [30,31]. We had to embed the questionnaire as a link in the SMS text message rather than have the questions directly in the message. Barriers to accessing the link could include patients not having a smartphone, although more than 90% of people in the United States have a smartphone [32]. Even if patients had a smartphone, patients may not have had access to the internet on the device to access the link. Another barrier to using the technology could have been that people were concerned about accessing a link on their phone, because of concerns about privacy, or were unsure of how to access the link.

Despite the low response rate and the limitations, this study provides important information for clinical care. This is the first study to assess patients' health-related social needs using SMS text messaging, and the response rate was lower than what has been seen in other SMS text messaging–based patient-reported outcomes studies [33,34]. Numerous national health care organizations have recommended that health systems address patients' unmet social needs as a routine part of clinical care, and CMS will require that all adult patients admitted to the hospital be screened for health-related social needs beginning in 2024 [6-9]. Many health systems are in the process of implementing different approaches to screen patients for social needs. At least in this population, simply sending out a SMS text message with the social risk questionnaire may not be effective as a single modality to assess all patients for health-related social needs. Health systems and clinics may need to implement multiple modalities, such as using SMS text messaging, the patient portal, or tablets in the clinic, to effectively screen all patients. If health systems are still interested in using SMS text messaging, they may also want to consider varying when and how the messages are sent.

There are several limitations to this study that should be acknowledged. First, although we included patients seen in 17 different primary care clinics, the clinics were all located in the same county and part of the same academic medical center so the results may not be generalizable to other populations. Second, all of the messages were sent at the same time and day for every patient. Varying when the message is sent (ie, time of the day and proximity to a clinic visit), may yield different results. Third, the message was sent based on the preferred language listed in the EHR. It is possible that the language listed was not correct. Fourth, we limited this study to patients with a diagnosis of hypertension. Future studies in other patient populations could find different results. Fifth, as in other studies screening patients for social risks in clinical care settings, individuals who have a primary care provider or clinic may be different than individuals who do not (ie, more likely to have health insurance) [35]. The results of the social risk questionnaire may not be representative of the surrounding community. Sixth, the SMS text message was unable to be delivered to 9 patients based on the phone number included in the EHR. As populations who have been socially and economically disadvantaged are more likely to have disruptions in their phone service, future research should evaluate the reasons why the message was unable to be delivered.

In this study, we found that only 17 (8.5%) of the 200 randomly selected patients completed a SMS text message–based health-related social needs questionnaire. Despite the negative results, this study provides important information for clinics considering implementing social needs screening. Further research is needed to understand how to most effectively and efficiently implement social needs screening in a patient-centered approach.

Acknowledgments

DP is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under grant (K23HL146902). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes

of Health. The funding organization has no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the study; and decision to submit the study for publication. DP reports personal fees from WellCare of North Carolina outside of the submitted work.

Data Availability

The data sets generated and analyzed during this study are not publicly available because they contain personal health information, but deidentified data sets are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data and questions included in text messages.
[DOCX File , 19 KB - cardio_v8i1e54530_app1.docx]

References

1. Berkowitz SA, Basu S. Unmet social needs and worse mental health after expiration of COVID-19 federal pandemic unemployment compensation. Health Aff (Millwood) 2021;40(3):426-434 [FREE Full text] [doi: 10.1377/hlthaff.2020.01990] [Medline: 33600235]
2. Berkowitz SA, Palakshappa D, Seligman HK, Hanmer J. Changes in food insecurity and changes in patient-reported outcomes: a nationally representative cohort study. J Gen Intern Med 2022;37(14):3638-3644 [FREE Full text] [doi: 10.1007/s11606-021-07293-4] [Medline: 34993877]
3. Berkowitz SA, Seligman HK, Meigs JB, Basu S. Food insecurity, healthcare utilization, and high cost: a longitudinal cohort study. Am J Manag Care 2018;24(9):399-404 [FREE Full text] [Medline: 30222918]
4. Seligman HK, Berkowitz SA. Aligning programs and policies to support food security and public health goals in the United States. Annu Rev Public Health 2019;40:319-337 [FREE Full text] [doi: 10.1146/annurev-publhealth-040218-044132] [Medline: 30444684]
5. Palakshappa D, Garg A, Peltz A, Wong CA, Cholera R, Berkowitz SA. Food insecurity was associated with greater family health care expenditures in the US, 2016-17. Health Aff (Millwood) 2023;42(1):44-52. [doi: 10.1377/hlthaff.2022.00414] [Medline: 36623217]
6. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Care Services, Committee on Integrating Social Needs Care into the Delivery of Health Care to Improve the Nation's Health. Integrating Social Care into the Delivery of Health Care: Moving Upstream to Improve the Nation's Health. Washington, DC: National Academies Press; 2019.
7. Accountable health communities model. Centers for Medicare and Medicaid Services. 2018. URL: https://www.cms.gov/priorities/innovation/innovation-models/ahcm [accessed 2024-01-23]
8. Parish W, Beil H, He F, D'Arcangelo N, Romaire M, Rojas-Smith L, et al. Health care impacts of resource navigation for health-related social needs in the accountable health communities model. Health Aff (Millwood) 2023;42(6):822-831. [doi: 10.1377/hlthaff.2022.01502] [Medline: 37196210]
9. Renaud J, McClellan SR, DePriest K, Witgert K, O'Connor S, Johnson KA, et al. Addressing health-related social needs via community resources: lessons from accountable health communities. Health Aff (Millwood) 2023;42(6):832-840. [doi: 10.1377/hlthaff.2022.01507] [Medline: 37196207]
10. Gottlieb L, Cottrell EK, Park B, Clark KD, Gold R, Fichtenberg C. Advancing social prescribing with implementation science. J Am Board Fam Med 2018;31(3):315-321 [FREE Full text] [doi: 10.3122/jabfm.2018.03.170249] [Medline: 29743213]
11. Gold R, Kaufmann J, Cottrell EK, Bunce A, Sheppler CR, Hoopes M, et al. Implementation support for a social risk screening and referral process in community health centers. NEJM Catal Innov Care Deliv 2023;4(4):CAT.23.0034 [FREE Full text] [doi: 10.1056/CAT.23.0034] [Medline: 37153938]
12. Palakshappa D, Benefield AJ, Furgurson KF, Harley MG, Bundy R, Moses A, et al. Feasibility of mobile technology to identify and address patients' unmet social needs in a primary care clinic. Popul Health Manag 2021;24(3):385-392 [FREE Full text] [doi: 10.1089/pop.2020.0059] [Medline: 32924796]
13. Gottlieb LM, Hessler D, Long D, Laves E, Burns AR, Amaya A, et al. Effects of social needs screening and in-person service navigation on child health: a randomized clinical trial. JAMA Pediatr 2016;170(11):e162521 [FREE Full text] [doi: 10.1001/jamapediatrics.2016.2521] [Medline: 27599265]
14. Steeves-Reece AL, Davis MM, Larson JH, Major-McDowall Z, King AE, Nicolaidis C, et al. Patients' willingness to accept social needs navigation after in-person versus remote screening. J Am Board Fam Med 2023;36(2):229-239 [FREE Full text] [doi: 10.3122/jabfm.2022.220259R1] [Medline: 36868871]

15. Gurewich D, Linsky AM, Harvey KL, Li M, Griesemer I, MacLaren RZ, et al. Relationship between unmet social needs and care access in a veteran cohort. J Gen Intern Med 2023;38(Suppl 3):841-848 [FREE Full text] [doi: 10.1007/s11606-023-08117-3] [Medline: 37340267]
16. Berkowitz SA, Seligman HK, Choudhry NK. Treat or eat: food insecurity, cost-related medication underuse, and unmet needs. Am J Med 2014;127(4):303-310.e3 [FREE Full text] [doi: 10.1016/j.amjmed.2014.01.002] [Medline: 24440543]
17. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022;60(1):95-103 [FREE Full text] [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
18. Ying X, Ehrhardt S. Pilot trial characteristics, postpilot design modifications, and feasibility of full-scale trials. JAMA Netw Open 2023;6(9):e2333642 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.33642] [Medline: 37707813]
19. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
20. RE-AIM. 2019. URL: http://www.re-aim.org [accessed 2019-12-30]
21. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
22. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. Am J Public Health 2013;103(6):e38-e46 [FREE Full text] [doi: 10.2105/AJPH.2013.301299] [Medline: 23597377]
23. Brown EM, Loomba V, De Marchis E, Aceves B, Molina M, Gottlieb LM. Patient and patient caregiver perspectives on social screening: a review of the literature. J Am Board Fam Med 2023;36(1):66-78 [FREE Full text] [doi: 10.3122/jabfm.2022.220211R1] [Medline: 36759136]
24. De Marchis EH, Alderwick H, Gottlieb LM. Do patients want help addressing social risks? J Am Board Fam Med 2020;33(2):170-175 [FREE Full text] [doi: 10.3122/jabfm.2020.02.190309] [Medline: 32179597]
25. Palakshappa D, Doupnik S, Vasan A, Khan S, Seifu L, Feudtner C, et al. Suburban families' experience with food insecurity screening in primary care practices. Pediatrics 2017;140(1):e20170320. [doi: 10.1542/peds.2017-0320] [Medline: 28634248]
26. Palakshappa D, Goodpasture M, Albertini L, Brown CL, Montez K, Skelton JA. Written versus verbal food insecurity screening in one primary care clinic. Acad Pediatr 2020;20(2):203-207 [FREE Full text] [doi: 10.1016/j.acap.2019.10.011] [Medline: 31629943]
27. Gottlieb L, Hessler D, Long D, Amaya A, Adler N. A randomized trial on screening for social determinants of health: the iScreen study. Pediatrics 2014;134(6):e1611-e1618. [doi: 10.1542/peds.2014-1439] [Medline: 25367545]
28. Cullen D, Woodford A, Fein J. Food for thought: a randomized trial of food insecurity screening in the emergency department. Acad Pediatr 2019;19(6):646-651 [FREE Full text] [doi: 10.1016/j.acap.2018.11.014] [Medline: 30639763]
29. Miller DP, Foley KL, Bundy R, Dharod A, Wright E, Dignan M, et al. Universal screening in primary care practices by self-administered tablet vs nursing staff. JAMA Netw Open 2022;5(3):e221480 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.1480] [Medline: 35258581]
30. Chan DYL, Lee SWH, Teh PL. Factors influencing technology use among low-income older adults: a systematic review. Heliyon 2023;9(9):e20111 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e20111] [Medline: 37809586]
31. Straub ET. Understanding technology adoption: theory and future directions for informal learning. Rev Educ Res 2017;79(2):625-649. [doi: 10.3102/0034654308325896]
32. Mobile fact sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2024-01-01]
33. Agarwal AK, Ali ZS, Shofer F, Xiong R, Hemmons J, Spencer E, et al. Testing digital methods of patient-reported outcomes data collection: prospective cluster randomized trial to test SMS text messaging and mobile surveys. JMIR Form Res 2022;6(3):e31894 [FREE Full text] [doi: 10.2196/31894] [Medline: 35298394]
34. Romeiser JL, Cavalcante J, Richman DC, Singh SM, Liang X, Pei A, et al. Comparing email, SMS, and concurrent mixed modes approaches to capture quality of recovery in the perioperative period: retrospective longitudinal cohort study. JMIR Form Res 2021;5(11):e25209 [FREE Full text] [doi: 10.2196/25209] [Medline: 34734827]
35. Garg A, Boynton-Jarrett R, Dworkin PH. Avoiding the unintended consequences of screening for social determinants of health. JAMA 2016;316(8):813-814. [doi: 10.1001/jama.2016.9282] [Medline: 27367226]

Abbreviations

AHWFB: Atrium Health Wake Forest Baptist Health
CMS: Centers for Medicare and Medicaid
EHR: electronic health record
RE-AIM: Reach, Effectiveness, Acceptability, Implementation, and Maintenance

Edited by A Mavragani; submitted 13.11.23; peer-reviewed by R Gold, L Gottlieb, A Sheon; comments to author 14.12.23; revised version received 03.01.24; accepted 16.01.24; published 13.02.24.

Please cite as:
Kormanis A, Quinones S, Obermiller C, Denizard-Thompson N, Palakshappa D
Feasibility of Using Text Messaging to Identify and Assist Patients With Hypertension With Health-Related Social Needs: Cross-Sectional Study
JMIR Cardio 2024;8:e54530
URL: https://cardio.jmir.org/2024/1/e54530
doi:10.2196/54530
PMID:38349714

Original Paper

Comparing the Efficacy of Targeted and Blast Portal Messaging in Message Opening Rate and Anticoagulation Initiation in Patients With Atrial Fibrillation in the Preventing Preventable Strokes Study II: Prospective Cohort Study

Alok Kapoor[1,2], MS, MD; Parth Patel[1], BS; Soumya Chennupati[1], BS; Daniel Mbusa[1], BS; Hammad Sadiq[1], BS; Sanjeev Rampam[1], BS; Robert Leung[1,2], MBA, MD; Megan Miller[3], BS; Kevin Rivera Vargas[3], BS; Patrick Fry[4], BS; Mary Martin Lowe[5], PhD; Christina Catalano[4], MBA; Charles Harrison[4], MD; John Nicholas Catanzaro[4], MD; Sybil Crawford[1], PhD; Anne Marie Smith[6], MBA

[1]University of Massachusetts Chan Medical School, Worcester, MA, United States

[2]University of Massachusetts Memorial Health Care, Worcester, MA, United States

[3]College of Pharmacy, University of Florida, Jacksonville, FL, United States

[4]College of Medicine, University of Florida, Jacksonville, FL, United States

[5]Learning Advisors LLC, Chicago, IL, United States

[6]Heart Rhythm Society, Washington, DC, United States

Corresponding Author:
Alok Kapoor, MS, MD
University of Massachusetts Chan Medical School
55 N Lake Ave
Worcester, MA, 01655
United States
Phone: 1 9178564538
Email: alok.kapoor@umassmemorial.org

Abstract

Background: The gap in anticoagulation use among patients with atrial fibrillation (AF) is a major public health threat. Inadequate patient education contributes to this gap. Patient portal–based messaging linked to educational materials may help bridge this gap, but the most effective messaging approach is unknown.

Objective: This study aims to compare the responsiveness of patients with AF to an AF or anticoagulation educational message between 2 portal messaging approaches: sending messages targeted at patients with upcoming outpatient appointments 1 week before their scheduled appointment (targeted) versus sending messages to all eligible patients in 1 blast, regardless of appointment scheduling status (blast), at 2 different health systems: the University of Massachusetts Chan Medical School (UMass) and the University of Florida College of Medicine-Jacksonville (UFL).

Methods: Using the 2 approaches, we sent patient portal messages to patients with AF and grouped patients by high-risk patients on anticoagulation (group 1), high-risk patients off anticoagulation (group 2), and low-risk patients who may become eligible for anticoagulation in the future (group 3). Risk was classified based on the congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age between 65 and 74 years, and sex category (CHA_2DS_2-VASc) score. The messages contained a link to the Upbeat website of the Heart Rhythm Society, which displays print and video materials about AF and anticoagulation. We then tracked message opening, review of the website, anticoagulation use, and administered patient surveys across messaging approaches and sites using Epic Systems (Epic Systems Corporation) electronic health record data and Google website traffic analytics. We then conducted chi-square tests to compare potential differences in the proportion of patients opening messages and other evaluation metrics, adjusting for potential confounders. All statistical analyses were performed in SAS (version 9.4; SAS Institute).

Results: We sent 1686 targeted messages and 1450 blast messages. Message opening was significantly higher with the targeted approach for patients on anticoagulation (723/1156, 62.5% vs 382/668, 57.2%; *P*=.005) and trended the same in patients off

anticoagulation; subsequent website reviews did not differ by messaging approach. More patients off anticoagulation at baseline started anticoagulation with the targeted approach than the blast approach (adjusted percentage 9.3% vs 2.1%; *P*<.001).

Conclusions: Patients were more responsive in terms of message opening and subsequent anticoagulation initiation with the targeted approach.

(JMIR Cardio 2024;8:e49590) doi:10.2196/49590

KEYWORDS

anticoagulants; atrial fibrillation; humans; outpatients; patient education as topic; patient portals

Introduction

About 6 million Americans have atrial fibrillation (AF), with 12 million projected by 2050 [1-3]. AF accounts for 15% of ischemic strokes, resulting in permanent disability in 60% of cases and death in up to 20% [4]. The main approach to stroke prevention is anticoagulation. Although guidelines [5] and evidence exist to guide providers in prescribing anticoagulation, only about 60% of eligible patients receive anticoagulation, leading to a projected annual excess stroke rate of 100,000 [6,7]. Low adherence to this guideline results from a combination of not initiating anticoagulation when indicated and discontinuing anticoagulation prematurely. This is particularly true in patients of minority race and ethnicity, where anticoagulation use is lower and stroke rates are higher [8-13].

There are multiple barriers to initiating and persisting with anticoagulation. Access to specialists, socioeconomic status, and health literacy each represent a barrier [8]. The advent of patient portals makes electronic messaging an attractive, low-cost method to educate patients and prepare them for visits with their anticoagulation providers. While the electronic health record (EHR) patient portal is increasingly being used in health care to improve patient education, engagement, and health outcomes, responsiveness to this methodology for anticoagulation use in patients with atrial fibrillation is unknown.

A recent review suggests that patient education about anticoagulation through a mobile device, such as a smartphone or tablet, increases patient knowledge levels, medication adherence, and satisfaction and is associated with improved clinical outcomes [14]. EHR-based programs have also been identified as a valuable method to improve warfarin therapy, a type of anticoagulation, self-management for pediatric patients with congenital heart diseases [15]. Evaluating patient responsiveness to different portal-based messaging methods can help identify the optimal use of EHR patient portal tools to best support patients in managing their AF and anticoagulation.

In this study, we compare patient responsiveness to 2 approaches to patient portal messaging with the goal of directing patients to the Upbeat website [16] of the Heart Rhythm Society, which contains print and video information about AF and anticoagulation.

Methods

Overview

We previously published the protocol for our paper, which covered the methods used at UMass to send patient messages [17]. We will briefly summarize the pertinent elements of the methods for that messaging campaign. We will also include additional details regarding the parallel messaging intervention at the UFL.

Study Design

We conducted a prospective cohort study. We sent patients a message through MyChart (Epic Systems Corporation), the patient portal associated with Epic Systems (Epic Systems Corporation) EHR, introducing the study and the purpose of communication (Multimedia Appendix 1). The message contained a link (unique to each site) to educational materials housed on a professional society web page—that is, the Upbeat website produced by the Heart Rhythm Society (HRS)—as well as a link to a survey soliciting feedback about the educational materials (Multimedia Appendix 2). Essentially, we created 2 unique websites, (1 for each site) but with the same content and layout (clone copies). In the first approach, at the University of Massachusetts Chan Medical School (UMass), we tested targeted messaging by sending messages to patients through MyChart 1 week before an appointment with a cardiology provider or primary care provider. In the second approach, at the University of Florida College of Medicine-Jacksonville (UFL), we tested a blast messaging approach of sending a message to all eligible patients independent of an appointment. At UMass, we facilitated the message-sending process with a bulk communication tool available through Epic Systems. At UFL, we sent messages manually.

Setting

We included the cardiology and primary care practices of the UMass Memorial Health System located in central Massachusetts, as well as the patients within UFL's ambulatory practices located in northern Florida and southern Georgia. Both sites used the Epic Systems EHR and the MyChart patient portal for the duration of the study. We sent messages to UMass patients from November 2021 until February 2022. At UFL, we sent all messages in November 2021.

Participants

We included patients aged 18 years or older with AF with active MyChart patient portal accounts and who had at least 1 office visit in the 12 months before the start of our messaging intervention in November 2021. At UMass, starting each workday from November 2021 to February 2022, we ran the Epic System's Reporting Workbench that identified patients based on their having an appointment (office or tele-visit type) with a primary or cardiology care provider scheduled to take place within a week. At UFL, from November 2021 to

December 2021, we identified patients based on a previously established registry of those patients with AF who had a visit with a primary or cardiology care provider in the previous year. At both sites, we grouped patients based on their anticoagulation status (eg, on or off anticoagulation) and congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, stroke, vascular disease, age 65-74 years, and sex scale (CHA_2DS_2-VASc) score. Specifically, group 1 included those at high risk (eg, CHA_2DS_2-VASc score of ≥2 for men and ≥3 for women) and currently on anticoagulation; group 2 included those at high risk and off anticoagulation; and group 3 included those at low risk (eg, CHA_2DS_2-VASc score <2 for men and <3 for women) and not on anticoagulation.

Outcomes, Variables, or Data Sources

Message Opening

We tracked message opening as the number of messages open divided by the number of messages sent. To identify messages, we relied on Epic Systems clarity structured query language–based coding. Specifically, we collected all messages received from individual patients and then filtered them by messages sent by the study coordinators. Study coordinators did not send messages for other purposes, allowing us to only isolate study-related messaging.

Website Review

Using Google Analytics (Google LLC), we tracked the number of unique page views as the value representing the total number of unique sessions. As patients may have received more than 1 message throughout the study (corresponding to 2 separate visits or in the case of canceled and rescheduled visits), we selected the number of messages sent as the denominator. We then calculated the percentage of messages resulting in a unique page view, with the number of unique page views as the numerator and compared this across sites. We also compared the "bounce rate" across sites, which represents the percentage of all sessions on a site in which users only viewed a single page. Google documentation [18] notes that bounce rates should be interpreted within the context of a specific website's purpose. Upbeat, the website our messages directed patients toward, has many links to educational resources regarding anticoagulation and AF. We consider navigation away from the landing page to indicate more patient engagement with these educational materials. Thus, having a lower bounce rate indicates higher engagement with the Upbeat website beyond the information presented on the landing page. Digital experience research indicates that a bounce rate of less than 40% is excellent [19], although the referenced source did not provide a specific bounce rate for health education websites, which may differ from other types of websites.

Survey-Based Outcomes

We compared survey responses across both messaging approaches and sites. For group 1 (high risk, on anticoagulation), the survey covered domains of discussions of personal stroke risk, history of anticoagulation use, and persistence. For group 2 (high risk, off anticoagulation), the survey covered discussions of personal stroke risk, the report by the patient of receiving a provider suggestion to take anticoagulation, and the reason for stopping anticoagulation for those with previous use. For group 3 (low risk, not on anticoagulation), the survey covered the likelihood of learning more about personal stroke risk, willingness to start anticoagulation, and reasons for anticoagulation hesitancy. We also asked all 3 groups of patients about their attitude toward the Upbeat website materials, including if the materials were understandable, useful, and something they would recommend to other patients. We collected responses on a 5-point Likert scale ranging from strongly disagree to strongly agree.

Anticoagulation Use After Messaging

We tracked anticoagulation use through medication and laboratory records from our EHR for the 3 months following the completion of our messaging program until May 2023. To be on anticoagulation, a patient had to have an active prescription for an anticoagulant updated at an office visit in the 12 months before the start of the messaging program in November 2021. Moreover, the prescription had to be consistent with a therapeutic dose to prevent strokes associated with AF. We assigned baseline status based on the presence or absence of an anticoagulation medication on the current medication list for a visit occurring in the 12 months before baseline. We also considered a patient to be on anticoagulation if they had an international normalized ratio value of 1.5 or higher recorded within 60 days of the date of the end of follow-up, following an example in the literature as well as the clinical threshold commonly observed to make decisions about surgery and anticoagulation reversal [20,21].

Independent Exposure

The independent exposure was the messaging approach used (targeted at UMass vs blast at UFL).

Other Exposures: Anticoagulation Outcome Only

We included stroke risk based on the CHA_2DS_2-VASc score, which is comprised of congestive heart failure, hypertension, age, diabetes, previous stroke, vascular disease, and gender. To adjust further for potential confounders of the association between anticoagulation use and message opening, we included demographics omitted in that score (ie, race, ethnicity, language preference, and primary insurance). Finally, we included chronic kidney disease and anemia. In general, we relied on the International Classification of Diseases, tenth edition codes for the presence of a comorbid condition. For chronic kidney disease, low platelet count, and anemia, we relied on laboratory data.

Analysis or Efforts to Address Bias, Study Size, and Statistical Methods

Although we did not calculate an effect size a priori for this study, in our previous work, we have typically attempted to find a 5% or greater increase in anticoagulation initiation. A 5% increase would correspond with the prevention of 5 strokes over 1 year at our sites and 5000 strokes per year in the United States. We derive these figures from a large national registry reporting stroke rates in patients with AF as well as other epidemiological studies [22,23].

Message Opening, Website Review, and Survey-Based Outcomes

We calculated a chi-square-based *P* value comparing proportions of patients, or in the case of website site review, unique sessions, across messaging approaches or sites.

Anticoagulation Outcome

Among patients opening portal messages, we compared anticoagulation across the 2 messaging approaches. Specifically, we compared anticoagulation use 3 months after completion of messaging with both approaches for patients in group 1 and then separately for those in group 2. For this outcome, we excluded patients who did not have information to calculate baseline anticoagulation status (ie, visits within the past 12 months where anticoagulation status would have been updated) [12]. We did not impute missing anticoagulation status given the number of missing values and the unclear randomness of missingness as suggested in guidance from the literature [24]. To determine the significance of the difference in the percentage of anticoagulation use across message approaches, we calculated a chi-square-based *P* value, comparing proportions of anticoagulation separately for group 1 and then again for group 2.

To address potential bias from the confounder of the difference in populations at the 2 different sites, we computed the adjusted percentage of patients on anticoagulation between messaging approaches. More specifically, we constructed a generalized logistic mixed model with anticoagulation status (on or off) as the dependent variable and messaging approach (targeted vs blast) as the independent variable. We also included a random effect for provider to account for potential clustering and several covariates to adjust for potential confounders of anticoagulation. Covariates included variables making it more likely to be on anticoagulation (eg, higher stroke risk expressed through the CHA_2DS_2-VASc score) as well as factors making it less likely to be on anticoagulation (eg, anemia, chronic kidney disease, and high BMI). We did this separately for groups 1 and 2.

We performed all calculations in SAS (version 9.4; SAS Institute). In Multimedia Appendix 3, we include the code used to conduct the analysis.

Ethical Considerations

At UMass, the institutional review board (IRB) approved this protocol with an implied consent process (ie, we argued consent would be implied by those choosing to review the website or answer our survey). We also provided patients with the opportunity to opt-out if they did not want us to use their information about message opening or anticoagulation use. At the time of analysis, all data were deidentified or anonymized by stripping real identifiers with a unique study identifier. All patients were informed before they provided implied informed consent. The UMass Chan IRB approved the waiver of documentation of written informed consent as the study was minimal-risk, appropriate confidentiality protections were to be exercised, and the waiver of consent would not adversely affect the rights and welfare of subjects. The authors designed the study and gathered and analyzed the data according to the Helsinki Declaration guidelines on human research. The research protocol used in this study was reviewed and approved by the UMass Chan IRB (H00021866). The authors did not use any form of AI in any portion of this study, including manuscript writing.

At UFL, the IRB exempted the study as quality improvement.

Results

Message Opening

We sent 1156 (UMass) and 668 (UFL) messages to group 1 patients, 438 and 632 messages to group 2 patients, and 92 and 150 messages to group 3 patients with the targeted and blast approaches, respectively. Cohort characteristics by group and messaging approach are described in Table 1.

Message opening was moderately high at both sites and across groups, with the highest opening rates in group 1 (723/1156, 62.5%) at UMass and group 3 (87/150, 57.3%) at UFL. Message opening in group 1 was significantly higher at UMass than at UFL (723/1156, 62.5% vs 382/668, 57.2%; *P*=.005). We did not find a statistically significant difference in message opening rates between the targeted (UMass) and blast (UFL) messaging approaches for group 2 (274/438, 62.6% vs 335/632, 53%; *P*=.09) and group 3 (52/92, 56.5% vs 86/150, 57.3%; *P*=.22).

Table 1. Key characteristics in patients receiving messages with targeted versus blast approach (percentages may not sum to 100 due to rounding).

Characteristics	Targeted messaging (University of Massachusetts)			Blast messaging (University of Florida)		
	Group 1 (high risk on anticoagulation; n=1156), n (%)	Group 2 (high risk off anticoagulation; n=438), n (%)	Group 3 (low risk; n=92), n (%)	Group 1 (high risk on anticoagulation; n=668), n (%)	Group 2 (high risk off anticoagulation; n=632), n (%)	Group 3 (low risk; n=150), n (%)
Age						
<65	272 (23.5)	96 (21.9)	15 (16.3)	197 (29.5)	204 (32.3)	103 (68.7)
65-74	374 (32.4)	140 (32)	44 (47.8)	256 (38.3)	192 (30.4)	40 (26.7)
≥75	501 (43.3)	196 (44.8)	31 (33.7)	213 (31.9)	221 (35)	6 (4)
Missing	9 (0.8)	6 (1.4)	2 (2.2)	2 (0.3)	15 (2.4)	1 (0.7)
Sex						
Female	436 (37.7)	186 (42.5)	33 (35.9)	291 (43.6)	292 (46.2)	37 (24.7)
Male	708 (61.2)	246 (56.2)	57 (62)	375 (56.1)	325 (51.4)	112 (74.7)
Missing	12 (1)	6 (1.4)	2 (2.2)	2 (0.3)	15 (2.4)	1 (0.7)
Race						
Black	18 (1.6)	4 (0.9)	0 (0)	181 (27.1)	121 (19.1)	225 (16.7)
Other	49 (4.2)	19 (4.3)	2 (2.2)	37 (5.5)	29 (4.6)	5 (3.3)
White	1080 (93.4)	414 (94.5)	90 (97.8)	447 (66.9)	463 (73.2)	119 (79.3)
Decline to answer, missing, or unknown	9 (0.8)	1 (0.2)	0 (0)	3 (0.4)	19 (3)	1 (0.7)
Ethnicity						
Hispanic or Latino	44 (3.8)	14 (3.2)	3 (3.3)	20 (3)	13 (2)	6 (4)
Not Hispanic or Latino	1089 (94.2)	420 (95.9)	88 (95.6)	644 (96.4)	597 (94.5)	142 (94.7)
Decline to Answer	22 (1.9)	4 (0.9)	1 (1.1)	1 (0.1)	2 (0.3)	0 (0)
Unknown or missing	1 (0.1)	0 (0)	0 (0)	3 (0.4)	20 (3.2)	2 (1.3)
Language preference						
English	1109 (95.9)	419 (95.7)	90 (97.8)	649 (97.2)	604 (95.6)	148 (98.7)
Not English	47 (4.1)	19 (4.3)	2 (2.2)	16 (2.4)	11 (1.7)	2 (1.3)
Unknown or missing	0 (0)	0 (0)	0 (0)	3 (0.4)	17 (2.7)	1 (0.7)

Website Review

Using Google Analytics, we observed that few patients reviewed the Upbeat website—80 and 76 unique page views (P=.56) at UMass and UFL, respectively. For those that did review the website, the number that interacted with only viewed a single page of the website, that is, the "bounce rate" across both sites was between 54% and 57%. While a bounce rate of 40% or less is generally considered good [19], the referenced source did not provide a specific bounce rate for health education websites, which may differ from other types of websites. Bounce rates are best understood in the context of a website's purpose and type. The average bounce rate for an informational website and landing pages tends to be higher than other website types [25], thus our findings indicate moderately high engagement with the Upbeat website.

The average session duration was shorter at UFL than at UMass (83 seconds vs 148 seconds). Although we can conduct a statistical test for the average session duration, Google Analytics did not provide the distribution of individual times for each unique page viewer (Table S1 in Multimedia Appendix 4 [19,26-29] for the remaining comparisons).

Survey-Based Outcomes

From Group 1, 93 and 59 patients answered our survey using the targeted and blast messaging approaches, respectively. There was not a significant difference in patient reports of discussion with their provider about stroke risk, the duration of current anticoagulation use, or the frequency of missing doses of anticoagulation. Notably, forgetfulness and other reasons (apart from costs, side effects, or lack of benefit) comprise the majority of reasons for forgetting doses across messaging approaches. Most patients in both the targeted vs blast messaging groups strongly agreed or agreed that the materials from the HRS were easy-to-understand (68/82, 83% vs 13/15, 87%), were useful (69/82, 84% vs 14/15, 93%), as well as something they would recommend (71/83, 85% vs 14/15, 93%) without any of the differences reaching statistical significance (Table S2 in Multimedia Appendix 4 [27] for details).

From Group 2, a total of 9/25 patients answered our survey using the targeted and blast messaging approaches, respectively. More patients in the UMass group had discussed their stroke risk with their physician at UMass (16/25, 64% vs 3/9, 33%; *P*=.04). Among the patients in the targeted approach, only 26% (6/25) reported concern about the risk of bleeding as a cause for stopping anticoagulation. Only 4 patients from the blast approach answered this item, limiting comparison. The majority of patients strongly agreed or agreed that the materials from the HRS were easy-to-understand and useful, as well as something they would recommend (also, comparisons were limited due to only 3 patients from the blast messaging group answering this item; Table S3 in Multimedia Appendix 4 [27]).

For group 3, we only had 2 responses from the targeted approach and 1 response from the blast approach and therefore did not conduct any further calculations or comparisons.

Anticoagulation

For this outcome, we excluded patients for whom we did not have information to calculate baseline anticoagulation status (ie, visits within the past 12 months where anticoagulation status would have been updated) [12]. Among the included patients on anticoagulation (group 1) who opened messages, there were 636 and 285 from the targeted messaging and blast messaging approaches, respectively. Most patients reported race as White, with 91.8% (584/636) and 84.2% (240/285) under the targeted messaging and blast messaging approaches, respectively (Table 2).

The percentage of patients from group 1 on anticoagulation did not differ between targeted versus blast messaging approaches. By contrast, 11.9% (21/176) versus 3% (3/100; *P*=.01) of patients from group 2 in targeted versus blast messaging were on anticoagulation at the end of follow-up (Table 3). This difference persisted after adjustment with an anticoagulation percentage of 9.3% versus 2.1% (*P*<.001; Table 3).

Table 2. Key characteristics of patients opening messages with a targeted versus blast approach (percentages may not sum to 100 due to rounding).

Characteristics	Targeted messaging (University of Massachusetts; n=636), n (%)	Blast messaging (University of Florida; n=285), n (%)
Age		
<65	60 (9.4)	9 (3.2)
65-74	200 (31.4)	80 (28.1)
≥75	376 (59.1)	196 (68.8)
Sex		
Female	248 (39)	117 (41.1)
Male	388 (61)	168 (58.9)
Race[a]		
Black	9 (1.4)	28 (9.8)
Hispanic	17 (2.7)	6 (2.1)
Other	17 (2.7)	11 (3.8)
White	584 (91.8)	240 (84.2)
Missing	9 (1.4)	0 (0)
Insurance		
Commercial	62 (9.7)	21 (7.4)
Medicare	522 (82.1)	242 (84.9)
Medicaid	20 (3.1)	0 (0)
Other or state health insurance exchange	32 (5)	7 (2.4)
Missing	0 (0)	15 (5.3)
Anemia[b]		
Yes	359 (56.4)	146 (51.2)
No	269 (42.3)	131 (46)
Unknown	8 (1.3)	8 (2.8)
Chronic kidney disease[c]		
Stage 1	162 (25.5)	63 (22.1)
Stage 2	195 (30.7)	95 (33.3)
Stage 3	219 (34.4)	101 (35.4)
Stage 4 or 5	60 (9.4)	22 (7.7)
Missing	0 (0)	4 (1.4)
BMI Group		
Morbid obesity[d]	51 (8)	15 (5.3)
Not morbidly obese	585 (92)	270 (94.7)
Anticoagulant use at baseline		
Yes	459 (72.2)	186 (65.3)
No	177 (27.8)	99 (34.7)
Antiplatelet use		
Yes	353 (55.5)	237 (83.2)
No	283 (44.5)	48 (16.8)

[a]Black includes Black of African American and multiracial, including Black or African American, Hispanic includes those individuals reporting Hispanic or Latino ethnicity, Asian includes White Hispanic. There were no individuals who reported Black race or Hispanic ethnicity. Other include Asian, Native American, Alaska Native, and others.

[b]Defined using established criteria, that is, hemoglobin <13 g/dL for male candidates and <12 g/dL for female candidates [30].

[c]Defined by established criteria [26] as a creatinine clearance calculated in mL/minute/1.73 m^2 units for each stage: >90 (stage 1), 60-80 (stage 2), 30-59 (stage 3), 15-29 (stage 4), and <15 (stage 5).

[d]Obesity indicates a BMI ≥40.0 kg/m^2, as defined by the World Health Organization [31].

Table 3. Unadjusted and adjusted percentages of anticoagulation for 2 messaging approaches or sites stratified by group.

	Unadjusted percentage on anticoagulation			Adjusted percentage on anticoagulation		
	Targeted messaging (UMass[a]), n (%)	Blast messaging (UFL[b]), n, (%)	*P* value[c]	Targeted messaging (UMass)[d], %	Blast messaging (UFL)[d], %	*P* value[e]
Group 1 (high risk, on anticoagulation at baseline)	438 (95.4)	179 (96.2)	.64	96.6	96.4	.52
Group 2 (high risk, off anticoagulation at baseline)	21 (11.9)	3 (3.0)	.01	9.3	2.1	<.001

[a]UMass: University of Massachusetts.

[b]UFL: University of Florida.

[c]Chi-square–based *P* value.

[d]Only percentages are shown.

[e]Value derived from generalized estimating equation adjusting for age, gender, BMI, patient race-ethnicity, insurance, CHA_2DS_2-VASc score, presence of anemia (ie, hemoglobin <13 g/dL for male candidates, <12 g/dL for female candidates, and level of chronic kidney disease).

Discussion

Principal Results

The message opening was significantly higher with the targeted approach for patients on anticoagulation. Subsequent website reviews were not different across approaches. Notably, 7.2% more patients off anticoagulation at baseline started anticoagulation with the targeted approach.

Comparison With Previous Work

Several published studies have examined the impact of portal messaging. Toscos et al [32] and Toscos et al [33] found that a multicomponent intervention that included sending portal messages led to higher AF knowledge and adherence in AF patients randomized to the intervention compared to controls. The authors only focused on patients who had already been prescribed anticoagulation and found higher rates of patient portal use, similar to what we found in terms of message opening in this patient group. Szilagyi et al [34] demonstrated a small increase in influenza vaccination rates (on the order of 1%-3%) for patients receiving portal messages versus those not receiving one, but the authors did not study the delivery of the message in targeted versus blast approaches as we did. By contrast, Halket et al [35] studied the use of targeted electronic portal messaging for hepatitis C screening. More specifically, they studied the effect of sending a patient portal message for patients having an appointment in the upcoming 6 months compared with sending this message to those without an upcoming visit. Compared to controls, they found that 10% more patients (59/227, 26% vs 52/318, 16.4%; *P*<.01) underwent screening with the targeted approach. The authors do not further report the optimal timing within 6 months for sending a message. Presumably closer to the time of the visit would achieve the best results.

The main implication of this study is that targeted messaging was more effective than blast messaging in achieving message opening for those on anticoagulation. There was a trend toward increased message opening among patients off anticoagulation (274/438, 62.6% for the targeted approach versus 335/632, 53% for the blast approach). This increased message opening may have explained some anticoagulation starts, but replication at other sites would be helpful in drawing firm conclusions. Given the low rates of website reviews it is unlikely that it contributed to anticoagulation starts and would not be valuable to include in future programs, at least in the way we delivered it (as a simple website link). Education provided directly in the message or within the health portal is likely to be more effective than requesting patients to review external websites.

There are other implications for our findings. The best approach to messaging patients should also factor in local resources. Our approach to sending targeted messaging required the daily execution of a workbench report and subsequent filtering and transmission of portal messages. In the future, we anticipate that we could automate the manual steps and link the messaging with portal messages sent to patients related to preparation for ambulatory visits. Blast messaging may be successful in other contexts, such as for anticipated health programs such as yearly vaccination campaigns, as previously demonstrated. Although we sent blast messaging manually, automation could likely replace the manual process that we undertook and would likely require less support from IT professionals to code compared with targeted messaging. The clinical context, along with the cost and availability of IT support, should therefore dictate the approach that institutions and providers take when determining how to deliver messages to their patients.

Limitations

We acknowledge several limitations of this proposed study. Most notably, we did not randomly allocate patients to messaging approaches. Each site pursued the approach of its preference. Thereby, baseline differences in populations and provider practice patterns may have explained some of our findings. This is especially true for the outcomes of message

opening and website review, where we did not have patient-level variables. For the anticoagulation outcome, we adjusted for known confounders of the use of anticoagulation, including demographics, stroke risk score, and bleeding risk factors (ie, anemia and chronic kidney disease). Many other factors, including other indications for anticoagulation, type of anticoagulant, baseline health literacy, and computer literacy, may, however, have been different across our sites. In addition, because the site of care dictated the receipt of one versus the other messaging approach and we had limited information about the reason for receiving care at one versus the other site, we did not pursue propensity or other causal inference modeling approaches. Other institutional-based programs may have explained the increase with the targeted approach. At the same time, we were not aware of any systemwide programs at our sites during the time that we conducted this study. Additionally, we did not specifically test the messages with patients in a human-centered design approach. A human-centered design approach has successfully overcome limitations in other messaging programs cited in the literature [32,36]. Oake et al [36] observed that an automated voice messaging response system for communicating anticoagulation testing and dosage schedules to patients led to improved anticoagulation monitoring. Another limitation was that we were not able to ascertain if patients read our message, only that our portal message was opened. In many cases, a family member will be opening the message. Website review and survey responses may be limited in the same way. Although education by proxy through a family member may lead to decisions to take anticoagulation or stay on it, we were not able to distinguish the discrete effect of direct versus proxy communication in the current study. Our results may also not generalize to non-White populations, which is significant given the lower adherence of non-Whites [8]. Lastly, it is important to note the impact of COVID-19 and the timing of the UFL messages on the project. UFL messages were sent out in December, with a follow-up in January. In addition to patients receiving holiday-related emails, COVID-19 was surging as well. It is unclear how these two variables may have impacted message opening.

Conclusion

In conclusion, message opening was significantly higher with the targeted approach for patients on anticoagulation. Subsequent website reviews were not different across approaches. More patients off anticoagulation at baseline started anticoagulation with the targeted approach. The best approach to messaging patients should also factor in local resources.

Acknowledgments

Funds were secured by the Heart and Rhythm Society through grants from the Bristol-Myers Squibb–Pfizer Alliance and Janssen. The funders had no role in the initiative's design, implementation, interpretation, or reporting.

Data Availability

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

Authors' Contributions

AK is responsible for the conceptualization, formal analysis, and project administration. AK, PP, SC, DM, HS, SR, RL, MML, KRV, PF, CC, CH, JNC, and SC were responsible for writing the manuscript. MML was responsible for project administration. SC was responsible for methodology and formal analysis. AM was responsible for project administration and funding acquisition.

Conflicts of Interest

AK has received research grant support from Pfizer through its Independent Grants for Learning and Change funding mechanism and from Bristol-Myers Squibb for Independent Medical Education Grants. More recently, he has received research grant support through a competitive process adjudicated and funded by the Bristol-Myers Squibb-Pfizer Alliance, which is formed by both Pfizer and Bristol-Myers Squibb. He has also been awarded a grant by Pfizer to examine conversations between patients and providers. HS, SR, and SC have also received research grant support from Bristol Meyers Squibb in the past 3 years (staff members or coinvestigators on the grants secured by AK). MML has received funding from and served on rheumatology and transthyretin amyloidosis research fellowship review panels for Pfizer. JNC has received research grant support from Bristol-Myers Squibb and Pfizer. All remaining authors have nothing to disclose.

Multimedia Appendix 1
Patient portal messages that were sent to patients in Groups 1, 2, and 3.
[PDF File (Adobe PDF File), 228 KB - cardio_v8i1e49590_app1.pdf]

Multimedia Appendix 2
Blank questionnaires that were completed by patients in Groups 1, 2, and 3.
[PDF File (Adobe PDF File), 102 KB - cardio_v8i1e49590_app2.pdf]

Multimedia Appendix 3
SAS Code for Analysis of Anticoagulation Initiation Following Two Patient Portal Messaging Programs.
[PDF File (Adobe PDF File), 161 KB - cardio_v8i1e49590_app3.pdf]

Multimedia Appendix 4
Supplemental Tables S1-S3.
[PDF File (Adobe PDF File), 165 KB - cardio_v8i1e49590_app4.pdf]

References

1. Ball J, Carrington MJ, McMurray JJV, Stewart S. Atrial fibrillation: profile and burden of an evolving epidemic in the 21st century. Int J Cardiol 2013;167(5):1807-1824. [doi: 10.1016/j.ijcard.2012.12.093] [Medline: 23380698]
2. Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) study. JAMA 2001;285(18):2370-2375 [FREE Full text] [doi: 10.1001/jama.285.18.2370] [Medline: 11343485]
3. Krijthe BP, Kunst A, Benjamin EJ, Lip GYH, Franco OH, Hofman A, et al. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. Eur Heart J 2013;34(35):2746-2751 [FREE Full text] [doi: 10.1093/eurheartj/eht280] [Medline: 23900699]
4. Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, et al. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. Circulation 2004;110(9):1042-1046 [FREE Full text] [doi: 10.1161/01.CIR.0000140263.20897.42] [Medline: 15313941]
5. Maisel WH, Stevenson LW. Atrial fibrillation in heart failure: epidemiology, pathophysiology, and rationale for therapy. Am J Cardiol 2003;91(6A):2D-8D. [doi: 10.1016/s0002-9149(02)03373-8] [Medline: 12670636]
6. Marzec LN, Wang J, Shah ND, Chan PS, Ting HH, Gosch KL, et al. Influence of direct oral anticoagulants on rates of oral anticoagulation for atrial fibrillation. J Am Coll Cardiol 2017;69(20):2475-2484 [FREE Full text] [doi: 10.1016/j.jacc.2017.03.540] [Medline: 28521884]
7. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991;22(8):983-988. [doi: 10.1161/01.str.22.8.983] [Medline: 1866765]
8. Essien UR, Holmes DN, Jackson LR, Fonarow GC, Mahaffey KW, Reiffel JA, et al. Association of race/ethnicity with oral anticoagulant use in patients with atrial fibrillation: findings from the outcomes registry for better informed treatment of atrial fibrillation II. JAMA Cardiol 2018;3(12):1174-1182 [FREE Full text] [doi: 10.1001/jamacardio.2018.3945] [Medline: 30484833]
9. Essien UR, Magnani JW, Gellad WF, Fine MJ, Hernandez I. Abstract 123: race, ethnicity, and sex-related differences in anticoagulation in patients newly diagnosed with atrial fibrillation. Circ: Cardiovascular Quality and Outcomes 2019;12(Suppl_1):A123. [doi: 10.1161/hcq.12.suppl_1.123]
10. Magnani JW, Norby FL, Agarwal SK, Soliman EZ, Chen LY, Loehr LR, et al. Racial differences in atrial fibrillation-related cardiovascular disease and mortality: the Atherosclerosis Risk in Communities (ARIC) study. JAMA Cardiol 2016;1(4):433-441 [FREE Full text] [doi: 10.1001/jamacardio.2016.1025] [Medline: 27438320]
11. Harrison MJ, Marshall J. Atrial fibrillation, TIAs and completed strokes. Stroke 1984;15(3):441-442. [doi: 10.1161/01.str.15.3.441] [Medline: 6729872]
12. Writing Group Members, January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, et al. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines and the Heart Rhythm Society. Heart Rhythm 2019;16(8):e66-e93 [FREE Full text] [doi: 10.1016/j.hrthm.2019.01.024] [Medline: 30703530]
13. Michigan anticoagulation quality improvement initiative. Anticoagulation Toolkit. 2017. URL: http://anticoagulationtoolkit.org/ [accessed 2020-06-02]
14. Jang I. A systematic review on mobile health applications' education program for patients taking oral anticoagulants. Int J Environ Res Public Health 2021;18(17):8902 [FREE Full text] [doi: 10.3390/ijerph18178902] [Medline: 34501492]
15. Jones S, Hislop JL, Gilmore H, Greenway A, Hibbard J, Monagle P, et al. Using an electronic medical record patient portal for warfarin self-management: empowering children and parents. Res Pract Thromb Haemost 2023;7(2):100066 [FREE Full text] [doi: 10.1016/j.rpth.2023.100066] [Medline: 36891277]
16. Keep the rhythm going. UpBeat. URL: https://upbeat.org/ [accessed 2024-01-11]
17. Fitzpatrick M, Sadiq H, Rampam S, Araia A, Miller M, Vargas KR, et al. Preventing preventable strokes: a study protocol to push guideline-driven atrial fibrillation patient education via patient portal. Cardiovasc Digit Health J 2022;3(5):241-246 [FREE Full text] [doi: 10.1016/j.cvdhj.2022.07.068] [Medline: 36310680]
18. Bounce rate. Google Analytics. URL: https://support.google.com/analytics/answer/1009409?hl=en#:~:text=About%20bounce%20rate&text=Bounce%20rate%20is%20single%2Dpage,request%20to%20the%20Analytics%20server [accessed 2023-12-28]

19. What is a good bounce rate? And what's a bad one? Fullstory Education Team. 2023. URL: https://www.fullstory.com/blog/what-is-a-good-bounce-rate/ [accessed 2023-12-28]
20. Sadiq H, Hoque L, Shi Q, Manning G, Crawford S, McManus D, et al. SUPPORT-AF III: supporting use of AC through provider prompting about oral anticoagulation therapy for AF. J Thromb Thrombolysis 2021;52(3):808-816. [doi: 10.1007/s11239-021-02420-8] [Medline: 33694097]
21. Spyropoulos AC, Douketis JD. How i treat anticoagulated patients undergoing an elective procedure or surgery. Blood 2012;120(15):2954-2962 [FREE Full text] [doi: 10.1182/blood-2012-06-415943] [Medline: 22932800]
22. Gladstone DJ, Bui E, Fang J, Laupacis A, Lindsay MP, Tu JV, et al. Potentially preventable strokes in high-risk patients with atrial fibrillation who are not adequately anticoagulated. Stroke 2009;40(1):235-240 [FREE Full text] [doi: 10.1161/STROKEAHA.108.516344] [Medline: 18757287]
23. Friberg L, Rosenqvist M, Lip GYH. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish atrial fibrillation cohort study. Eur Heart J 2012;33(12):1500-1510 [FREE Full text] [doi: 10.1093/eurheartj/ehr488] [Medline: 22246443]
24. Lee JH, Huber JC. Evaluation of multiple imputation with large proportions of missing data: how much is too much? Iran J Public Health 2021;50(7):1372-1380 [FREE Full text] [doi: 10.18502/ijph.v50i7.6626] [Medline: 34568175]
25. Bounce rate benchmarks: what's a good bounce rate, anyway? CXL. URL: https://cxl.com/guides/bounce-rate/benchmarks/ [accessed 2023-10-23]
26. The difference between Google ads clicks, and sessions, users, entrances, pageviews, and unique pageviews in analytics. Google Analytics. URL: https://support.google.com/analytics/answer/1257084?hl=en#zippy=%2Cin-this-article [accessed 2023-02-13]
27. Atrial fibrillation. Heart Rhythm Society. URL: https://upbeat.org/heart-rhythm-disorders/atrial-fibrillation-afib [accessed 2023-02-13]
28. Alliance for aging research. Living with atrial fibrillation. YouTube. 2014. URL: https://www.youtube.com/watch?v=zj3rLviG1YA [accessed 2023-02-13]
29. Understanding the diagnosis of atrial fibrillation and how to prevent stroke. Heart Rhythm Society. URL: https://upbeat.org/documents/UnderstandAFibDiagnosis/download [accessed 2023-02-14]
30. Fang MC, Go AS, Chang Y, Borowsky LH, Pomernacki NK, Udaltsova N, et al. A new risk scheme to predict warfarin-associated hemorrhage: the ATRIA (Anticoagulation and Risk Factors in Atrial Fibrillation) study. J Am Coll Cardiol 2011;58(4):395-401 [FREE Full text] [doi: 10.1016/j.jacc.2011.03.031] [Medline: 21757117]
31. Mean Body Mass Index (BMI). World Health Organization. 2017. URL: https://www.who.int/gho/ncd/risk_factors/bmi_text/en/ [accessed 2020-01-10]
32. Toscos T, Coupe A, Wagner S, Ahmed R, Roebuck A, Flanagan M, et al. Engaging patients in atrial fibrillation management via digital health technology: the impact of tailored messaging. J Innov Card Rhythm Manag 2020;11(8):4209-4217 [FREE Full text] [doi: 10.19102/icrm.2020.110802] [Medline: 32874747]
33. Toscos T, Drouin M, Pater JA, Flanagan M, Wagner S, Coupe A, et al. Medication adherence for atrial fibrillation patients: triangulating measures from a smart pill bottle, e-prescribing software, and patient communication through the electronic health record. JAMIA Open 2020;3(2):233-242 [FREE Full text] [doi: 10.1093/jamiaopen/ooaa007] [Medline: 32734164]
34. Szilagyi PG, Albertin C, Casillas A, Valderrama R, Duru OK, Ong MK, et al. Effect of patient portal reminders sent by a health care system on influenza vaccination rates: a randomized clinical trial. JAMA Intern Med 2020;180(7):962-970 [FREE Full text] [doi: 10.1001/jamainternmed.2020.1602] [Medline: 32421168]
35. Halket D, Dang J, Phadke A, Jayasekera C, Kim WR, Kwo P, et al. Targeted electronic patient portal messaging increases hepatitis C virus screening in primary care: a randomized study. J Gen Intern Med 2022;37(13):3318-3324 [FREE Full text] [doi: 10.1007/s11606-022-07460-1] [Medline: 35230622]
36. Oake N, van Walraven C, Rodger MA, Forster AJ. Effect of an interactive voice response system on oral anticoagulant management. CMAJ 2009;180(9):927-933 [FREE Full text] [doi: 10.1503/cmaj.081659] [Medline: 19398739]

Abbreviations

AF: atrial fibrillation
CHA2DS2-VASc: congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age 65-74 years, sex scale
EHR: electronic health record
HRS: Heart Rhythm Society
IRB: institutional review board
UFL: University of Florida College of Medicine-Jacksonville
UMass: University of Massachusetts Chan Medical School

Edited by A Mavragani; submitted 08.06.23; peer-reviewed by M Nomali, M Wright; comments to author 28.09.23; revised version received 24.10.23; accepted 24.11.23; published 24.01.24.

Please cite as:
Kapoor A, Patel P, Chennupati S, Mbusa D, Sadiq H, Rampam S, Leung R, Miller M, Vargas KR, Fry P, Lowe MM, Catalano C, Harrison C, Catanzaro JN, Crawford S, Smith AM
Comparing the Efficacy of Targeted and Blast Portal Messaging in Message Opening Rate and Anticoagulation Initiation in Patients With Atrial Fibrillation in the Preventing Preventable Strokes Study II: Prospective Cohort Study
JMIR Cardio 2024;8:e49590
URL: https://cardio.jmir.org/2024/1/e49590
doi:10.2196/49590
PMID:38265849

Original Paper

Physical Activity, Heart Rate Variability, and Ventricular Arrhythmia During the COVID-19 Lockdown: Retrospective Cohort Study

Sikander Z Texiwala[1], MD; Russell J de Souza[2,3], RD, ScD; Suzette Turner[1,4], BScN; Sheldon M Singh[1,5], MD

[1]Schulich Heart Center, Sunnybrook Health Sciences, Toronto, ON, Canada

[2]Department of Health Research Methods, Evidence, and Impact, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[3]Population Health Research Institute, Hamilton Health Sciences Corporation, Hamilton, ON, Canada

[4]Lawrence Bloomberg Faculty of Nursing, University of Toronto, ON, Canada

[5]Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Sheldon M Singh, MD
Schulich Heart Center
Sunnybrook Health Sciences
Room A222
2075 Bayview Ave
Toronto, ON, M4N 3M5
Canada
Phone: 1 416 480 6100 ext 86359
Fax: 1 416 480 5707
Email: sheldon.singh@sunnybrook.ca

Abstract

Background: Ventricular arrhythmias (VAs) increase with stress and national disasters. Prior research has reported that VA did not increase during the onset of the COVID-19 lockdown in March 2020, and the mechanism for this is unknown.

Objective: This study aimed to report the presence of VA and changes in 2 factors associated with VA (physical activity and heart rate variability [HRV]) at the onset of COVID-19 lockdown measures in Ontario, Canada.

Methods: Patients with implantable cardioverter defibrillator (ICD) followed at a regional cardiac center in Ontario, Canada with data available for both HRV and physical activity between March 1 and 31, 2020, were included. HRV, physical activity, and the presence of VA were determined during the pre- (March 1-10, 2020) and immediate postlockdown (March 11-31) period. When available, these data were determined for the same period in 2019.

Results: In total, 68 patients had complete data for 2020, and 40 patients had complete data for 2019. Three (7.5%) patients had VA in March 2019, whereas none had VA in March 2020 (P=.048). Physical activity was reduced during the postlockdown period (mean 2.3, SD 1.6 hours vs mean 2.1, SD 1.6 hours; P=.003). HRV was unchanged during the pre- and postlockdown period (mean 91, SD 30 ms vs mean 92, SD 28 ms; P=.84).

Conclusions: VA was infrequent during the COVID-19 pandemic. A reduction in physical activity with lockdown maneuvers may explain this observation.

(JMIR Cardio 2024;8:e51399) doi:10.2196/51399

KEYWORDS

implantable cardioverter defibrillator; heart rate variability; physical activity; lockdown; ICD; ventricular arrhythmias; defibrillator; implementation

Introduction

Increased ventricular arrhythmias (VAs) have been reported with acts of terror and environmental disasters [1,2]. The onset of the COVID-19 pandemic was associated with increased levels of stress [3]. Although one may have anticipated an increased rate of VA during this time, this was not borne in North America and Europe [4,5]. A mechanism to explain this has not been elucidated.

Patients with implantable cardioverter defibrillators (ICDs) are at risk for VA. ICDs contain sensors that can quantify the physical activity of a patient who has an ICD implanted. Acute

increases in physical activity can increase the risk of VA [6]. ICDs monitor the changes in the patient's heart rate. Heart rate variability (HRV) summarizes the beat-to-beat changes in heart rate and reflects the balance between the sympathetic and parasympathetic nervous system. A reduction in HRV can occur during times of stress due to increased sympathetic activation. A reduction in HRV predicts VA [7]. Assessing the changes in physical activity and HRV may provide insight into the lack of increased VA observed during the onset of the COVID-19 lockdown.

Herein, we report the changes in physical activity and HRV in patients with ICD during the COVID-19 lockdown of March 2020 in Ontario, Canada.

Methods

Study Cohort

Sunnybrook Hospital is a regional cardiac center in Ontario, Canada. The Sunnybrook ICD clinical electronic medical record database (Paceart Optima System, version 1.8.269.0; Medtronic) was searched to identify all actively followed patients with ICD with data on physical activity and HRV in March 2020. As this was a retrospective observational study, patients were not actively recruited rather all patients with the available data were included in this retrospective cohort study.

Study Periods

Our study focused on the first month of the COVID-19 pandemic (March 2020). March 1-10, 2020, was designated as the prelockdown period and March 11-31, 2020, as the lockdown period. During the latter period, the World Health Organization declared the COVID-19 outbreak a pandemic (March 11), with a subsequent crash in North American stock markets (March 12) and a declaration of a state of emergency in the United States (March 13) and Ontario (March 17).

VA, Physical Activity, and Heart Rate Variability

ICDs record all VA. The presence of VA requiring an ICD therapy during the study period was documented. Patient physical activity is recorded when a patient moves at a rate above a minimum threshold of 2 miles per hour. ICDs quantify the amount of time spent moving above this rate.

ICD algorithms determine HRV as the SD of the average sinus intervals over 5 minutes, averaged over 24 hours (288 periods). This time domain approach to determine HRV provides the best prognostic information [7]. As the knowledge of atrial activity is necessary to determine the HRV, it cannot be determined in patients who do not have a dual chamber ICD (ie, a device with an atrial lead). Furthermore, HRV cannot be determined in the presence of inherently irregular arrhythmia such as atrial fibrillation. As such, patients with a single chamber ICD and a history of atrial fibrillation were excluded from the study.

Physical activity and HRV during the study period were extracted using an open-source software tool (WebPlotDigitizer, version 4.4). To provide an estimate of the change in physical activity and HRV between the 2 study periods, the extracted daily values for physical activity and HRV were averaged over the pre- and lockdown periods. Where available, VA, physical activity, and HRV were obtained from March 1 to 31, 2019, to act as a control.

Statistical Analysis

Participant characteristics are presented as mean (SD) or counts (%). Chi-square testing was used to determine the differences in the percentage of patients experiencing VA during the study periods. Two-tailed paired *t* tests and analysis of covariance were used to compare HRV and physical activity between the 2 study periods.

Statistical analyses were performed using the statistical analysis system statistical software package (version 9.4; SAS Institute Inc). *P* values <.05 were considered statistically significant.

Ethical Considerations

The Sunnybrook Hospital research ethics board approved the study (study 1632). The requirement for patient consent was waived by the research ethics board. Data were collected in anonymously. There was no patient compensation for participation in this study.

Results

Of the 650 actively followed ICD patients, 485 did not have data on both physical activity and HRV, and 97 patients did not have follow-up during the COVID-19 pandemic. The final cohort comprised of 68 patients, 40 of whom had HRV and physical activity data for 2021 and 2020.

The average age of the cohort was 70 (SD 11) years and predominantly male (n=52, 77%). Half of the participants (n=34, 50%) had coronary artery disease, 40% (n=27) had ventricular tachycardia, and 37% (n=25) had a cardiac resynchronization device. Beta-blockers and antiarrhythmic drugs were used by 77% (n=52) and 24% (n=16) of the cohort, respectively.

No patient had a VA during the 2020 study period. Three (7.5%) patients experienced VA in 2019, all between March 11 and 30. Thus, there were fewer VA events in the lockdown period of 2020 compared to the equivalent time in 2019 (n=0, 0% vs n=3, 7.5%; *P*=.048).

Activity was reduced by approximately 12 minutes during the lockdown period of 2020 compared to the prelockdown period of 2020 (mean 2.1, SD 1.6 hours vs mean 2.3, SD 1.6 hours; *P*=.003). There was no difference in the average activity in the prelockdown period in 2020 compared to the equivalent dates in 2019 (mean 2.3, SD 1.6 hours vs mean 2.5, SD 1.7 hours; *P*=.06).

HRV was unchanged between the 2020 prelockdown and lockdown periods (mean 91, SD 30 ms vs mean 92, SD 28 ms; *P*=.84). HRV was similar in 2020 and 2019 (prelockdown: mean 89, SD 28 ms vs mean 90, SD 30 ms; *P*=.70 and lockdown: mean 86, SD 26 ms vs mean 90, SD 25 ms; *P*=.30).

Discussion

Principal Findings

This work supports prior publications highlighting a lack of increased VA with the onset of the COVID-19 pandemic. Our

research is hypothesis generating, which provides a possible mechanism to explain the lack of increased VA observed at the onset of the COVID-19 pandemic. It is speculated that a reduction in physical activity with lockdown maneuvers may have reduced the frequency of VA.

Lockdown maneuvers resulted in stay-at-home orders, closing of gyms or shopping centers, and working from home. These maneuvers, which were similar in Ontario and other jurisdictions, effectively reduced the physical activity of all individuals during this time period. Although small (~12 minutes), it is possible that the reduction in physical activity may have played a role in mitigating arrhythmic risk. For context, a reduction of 12 minutes is at a minimum equivalent to a reduction in walking 0.4 miles or 1000 steps a day. As we are not able to quantify the intensity of the activity, it is possible this reduction in activity could have been higher.

HRV is a marker of autonomic tone and a predictor of VA. We observed no clinically important or statistically significant change in HRV. This finding seems counterintuitive, given the reports of increased distress with the onset of the lockdown maneuvers [3]. We hypothesize that a number of factors may have mitigated additional reductions in HRV in this population. First, the use of beta-blockers was high (n=52, 77%) in this population. Prior work has demonstrated that beta-blockers can preserve autonomic balance in the setting of mentally stressful events [8]. Second, patients with ICD already have a high level of circulating catecholamines (evidenced by the depressed HRV even prior to the COVID-19 pandemic). The additional influence of external psychological stresses with the COVID-19 pandemic may not impact overall autonomic tone. Finally, we speculate that, unlike singular unexpected catastrophic events [8], the anticipation of lockdown measures may have lessened this psychological stress. This finding highlights the variable impact of different catastrophic events on the risk of VA.

Limitations to this work include the fact the data were derived from a single center with a relatively small number of patients. Second, the large number of exclusions may have resulted in a highly selected population limiting the applicability to other populations. Third, we limited our assessment to the early part of the pandemic, given the homogenous lockdown interventions and limited impact of lack of access to care during this early time. It is unclear whether these findings would persist into different waves of the pandemic. Finally, the findings were from Ontario, Canada, and may not apply to other jurisdictions with more severe COVID-19 outbreaks and interventions. The primary strength of this work is the precise measure of VA, physical activity, and HRV.

Conclusions

Physical activity was reduced in patients with ICD during the COVID-19 lockdown. Our observations may provide a possible mechanistic insight into lack of increased VA in patients with ICD during the COVID-19 pandemic. We suggest future work in larger patient populations and other jurisdictions to confirm our findings. Given the long-term benefits of physical activity, we also suggest future work by public health agencies to ensure the observed decline in physical activity at the onset of the COVID-19 pandemic is not sustained.

Acknowledgments

This work was supported by a generous donation from the Horgan Family and Sunnybrook Hospital Foundation.

Data Availability

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

Authors' Contributions

All authors contributed to the design of the work, acquisition, analysis, and interpretation of the data. SZT and SMS drafted the manuscript. All authors critically revised the manuscript for important intellectual content. All authors approved the final version of the manuscript. All authors are accountable for the accuracy and integrity of the work. No artificial intelligence assistive tools were used to generate any portions of this work.

Conflicts of Interest

None declared.

References

1. Steinberg JS, Arshad A, Kowalski M, Kukar A, Suma V, Vloka M, et al. Increased incidence of life-threatening ventricular arrhythmias in implantable defibrillator patients after the World Trade Center attack. J Am Coll Cardiol 2004 Sep 15;44(6):1261-1264 [FREE Full text] [doi: 10.1016/j.jacc.2004.06.032] [Medline: 15364329]
2. Ranjbar F, Akbarzadeh F, Kazemi B, Ranjbar A, Namin SS, Sadeghi-Bazargani H. Increased likelihood of arrhythmic events associated with increased anxiety in patients with implanted cardiac defibrillators after the ahar-varzegan earthquake in East Azarbaijan, 2012. Bull Emerg Trauma 2016;4(4):202-210 [FREE Full text] [Medline: 27878125]
3. Manchia M, Gathier AW, Yapici-Eser H, Schmidt MV, de Quervain D, van Amelsvoort T, et al. The impact of the prolonged COVID-19 pandemic on stress resilience and mental health: a critical review across waves. Eur Neuropsychopharmacol 2022;55:22-83 [FREE Full text] [doi: 10.1016/j.euroneuro.2021.10.864] [Medline: 34818601]

4. O'Shea CJ, Thomas G, Middeldorp ME, Harper C, Elliott AD, Ray N, et al. Ventricular arrhythmia burden during the coronavirus disease 2019 (COVID-19) pandemic. Eur Heart J 2021 Feb 01;42(5):520-528 [FREE Full text] [doi: 10.1093/eurheartj/ehaa893] [Medline: 33321517]
5. Circulation. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation 1996 Mar 01;93(5):1043-1065. [Medline: 8598068]
6. Lampert R, Joska T, Burg MM, Batsford WP, McPherson CA, Jain D. Emotional and physical precipitants of ventricular arrhythmia. Circulation 2002;106(14):1800-1805 [FREE Full text] [doi: 10.1161/01.cir.0000031733.51374.c1] [Medline: 12356633]
7. Task Force for the European society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation 1996 Mar 01;93(5):1043-1065. [Medline: 8598068]
8. Huang JL, Chiou CW, Ting CT, Chen YT, Chen SA. Sudden changes in heart rate variability during the 1999 Taiwan earthquake. Am J Cardiol 2001;87(2):245-248, A9. [doi: 10.1016/s0002-9149(00)01331-x] [Medline: 11152854]

Abbreviations

HRV: heart rate variability
ICD: implantable cardioverter defibrillator
VA: ventricular arrhythmia

Edited by A Mavragani; submitted 30.07.23; peer-reviewed by C Darling, Y Chu; comments to author 03.11.23; revised version received 18.11.23; accepted 26.12.23; published 05.02.24.

Please cite as:
Texiwala SZ, de Souza RJ, Turner S, Singh SM
Physical Activity, Heart Rate Variability, and Ventricular Arrhythmia During the COVID-19 Lockdown: Retrospective Cohort Study
JMIR Cardio 2024;8:e51399
URL: https://cardio.jmir.org/2024/1/e51399
doi:10.2196/51399
PMID:38315512

Original Paper

User Engagement, Acceptability, and Clinical Markers in a Digital Health Program for Nonalcoholic Fatty Liver Disease: Prospective, Single-Arm Feasibility Study

Sigridur Björnsdottir[1], MD, PhD; Hildigunnur Ulfsdottir[2], MPH, MD; Elias Freyr Gudmundsson[2], BSc, MSc; Kolbrun Sveinsdottir[2], MSc; Ari Pall Isberg[2], MSc; Bartosz Dobies[2], MSc; Gudlaug Erla Akerlie Magnusdottir[2], BSc; Thrudur Gunnarsdottir[2], PhD; Tekla Karlsdottir[2], MD; Gudlaug Bjornsdottir[3,4], MSc; Sigurdur Sigurdsson[3,4], PhD; Saemundur Oddsson[2], MD; Vilmundur Gudnason[3,4], MD, PhD

[1]Department of Endocrinology, Metabolism and Diabetes, Karolinska Institutet, Stockholm, Sweden

[2]Sidekick Health, Kopavogur, Iceland

[3]Icelandic Heart Association, Kopavogur, Iceland

[4]School of Health Sciences, Faculty of Medicine, University of Iceland, Reykjavik, Iceland

Corresponding Author:
Sigridur Björnsdottir, MD, PhD
Department of Endocrinology, Metabolism and Diabetes
Karolinska Institutet
Solnavägen 1
Stockholm, 171 77
Sweden
Phone: 46 8 524 800 00
Email: sigridur.bjornsdottir@ki.se

Abstract

Background: Nonalcoholic fatty liver disease (NAFLD) has become the most common chronic liver disease in the world. Common comorbidities are central obesity, type 2 diabetes mellitus, dyslipidemia, and metabolic syndrome. Cardiovascular disease is the most common cause of death among people with NAFLD, and lifestyle changes can improve health outcomes.

Objective: This study aims to explore the acceptability of a digital health program in terms of engagement, retention, and user satisfaction in addition to exploring changes in clinical outcomes, such as weight, cardiometabolic risk factors, and health-related quality of life.

Methods: We conducted a prospective, open-label, single-arm, 12-week study including 38 individuals with either a BMI >30, metabolic syndrome, or type 2 diabetes mellitus and NAFLD screened by FibroScan. An NAFLD-specific digital health program focused on disease education, lowering carbohydrates in the diet, food logging, increasing activity level, reducing stress, and healthy lifestyle coaching was offered to participants. The coach provided weekly feedback on food logs and other in-app activities and opportunities for participants to ask questions. The coaching was active throughout the 12-week intervention period. The primary outcome was feasibility and acceptability of the 12-week program, assessed through patient engagement, retention, and satisfaction with the program. Secondary outcomes included changes in weight, liver fat, body composition, and other cardiometabolic clinical parameters at baseline and 12 weeks.

Results: In total, 38 individuals were included in the study (median age 59.5, IQR 46.3-68.8 years; n=23, 61% female). Overall, 34 (89%) participants completed the program and 29 (76%) were active during the 12-week program period. The median satisfaction score was 6.3 (IQR 5.8-6.7) of 7. Mean weight loss was 3.5 (SD 3.7) kg (P<.001) or 3.2% (SD 3.4%), with a 2.2 (SD 2.7) kg reduction in fat mass (P<.001). Relative liver fat reduction was 19.4% (SD 23.9%). Systolic blood pressure was reduced by 6.0 (SD 13.5) mmHg (P=.009). The median reduction was 0.14 (IQR 0-0.47) mmol/L for triglyceride levels (P=.003), 3.2 (IQR 0.0-5.4) µU/ml for serum insulin (s-insulin) levels (P=.003), and 0.5 (IQR –0.7 to 3.8) mmol/mol for hemoglobin A_{1c} (HbA_{1c}) levels (P=.03). Participants who were highly engaged (ie, who used the app at least 5 days per week) had greater weight loss and liver fat reduction.

Conclusions: The 12-week-long digital health program was feasible for individuals with NAFLD, receiving high user engagement, retention, and satisfaction. Improved liver-specific and cardiometabolic health was observed, and more engaged participants

showed greater improvements. This digital health program could provide a new tool to improve health outcomes in people with NAFLD.

Trial Registration: Clinicaltrials.gov NCT05426382; https://clinicaltrials.gov/study/NCT05426382

(JMIR Cardio 2024;8:e52576) doi:10.2196/52576

KEYWORDS

digital health program; nonalcoholic fatty liver disease; NAFLD; cardiometabolic health; digital therapeutics; liver; chronic; hepatic; cardiometabolic; cardiovascular; cardiology; weight; acceptability; digital health; metabolic syndrome; diabetic; diabetes; diabetics; type 2; BMI; lifestyle; exercise; physical activity; coaching; diet; dietary; nutrition; nutritional; patient education; coach; feasibility; fat; body composition

Introduction

Nonalcoholic fatty liver disease (NAFLD) is the most common chronic liver disease in the world [1]. NAFLD is defined as >5% fat in the liver (steatosis) among people who drink moderate amounts or no alcohol and have no other chronic liver diseases [2]. NAFLD reflects a spectrum of liver pathologies, ranging from simple steatosis to a more severe condition called nonalcoholic steatohepatitis (NASH), which includes inflammation and potential scarring of the liver [3]. The major comorbidities associated with NAFLD are central obesity, type 2 diabetes mellitus (DM), dyslipidemia, and metabolic syndrome [4]. The global prevalence of NAFLD is estimated to be 25% in the general population and the rising prevalence of NAFLD parallels that of obesity and type 2 DM, since NAFLD is a comorbidity in an estimated 55% of people with type 2 DM and in up to 80% of people with obesity [4-6]. Studies have shown that around 20% to 30% of people with NAFLD progress to NASH, with its consequent risks of liver scarring, cirrhosis, end-stage liver disease, and hepatocellular carcinoma [7]. Furthermore, NAFLD and NASH are associated with cardiovascular diseases, type 2 DM, and chronic kidney disease and pose a large burden on health care systems [8-10].

Growing evidence supports a common pathophysiological mechanism between metabolic syndrome and NAFLD and NASH, which often involves insulin resistance and dysfunctional adipose tissue [11]. Currently, no pharmacological treatment is approved for NAFLD or NASH, and, according to treatment guidelines, first-line therapy should focus on lifestyle improvement with the aim of 5% to 10% weight loss [12,13]. However, reaching these goals is often difficult, and there is a need to continue exploring optimal treatment modalities for individuals with NAFLD or NASH [14].

Sidekick Health, an Icelandic digital therapeutic company, has developed a digital health program (Sidekick-241 or SK-241) specifically designed for people with metabolic conditions and NAFLD. The 12-week program is delivered through a mobile app and aims to improve lifestyle and health outcomes by focusing on improving diet, increasing activity levels, and reducing stress through behavior change. In this prospective study, we evaluated the feasibility and potential clinical impact of the 12-week digital health program on liver and cardiometabolic health in individuals with metabolic conditions and NAFLD.

Methods

Trial Design

This was an open-label, single-arm, prospective study conducted between June and September 2022 in Iceland. The study included a 12-week digital health program delivered through the Sidekick app. Screening and preprogram and postprogram clinical assessments were carried out at the Icelandic Heart Association.

Participants

In total, 38 individuals aged between 18 and 80 years from an ongoing population-based cohort study (The REFINE-Reykjavik Study) at The Icelandic Heart Association and individuals followed at an endocrine outpatient clinic (the Reykjavik Heart Center) were invited to participate in the study [15]. People with at least one of the following risk factors were invited to participate: BMI >30, metabolic syndrome, or type 2 DM. Individuals with type 2 DM were only included if they were on a stable dose of antidiabetes medication for the last 90 days before screening. Eligible individuals had to have the capacity to give informed consent, understand verbal and written Icelandic, own and know how to operate a smartphone, and be willing and able to comply with the study program, all scheduled visits, and procedures.

The exclusion criteria were as follows: insulin use; known or self-reported cirrhosis; alcohol consumption over 14 units/week for men or over 7 units/week for women; self-reported hepatitis B, hepatitis C, human immunodeficiency virus, or autoimmune hepatitis; vitamin E intake of >400 IU/day unless stable for 12 weeks prior to baseline; taking medications associated with liver steatosis, such as steroids, methotrexate, tamoxifen, amiodarone, tetracycline, or valproic acid; self-reported pregnancy; participation in a weight loss program; or history of or any existing medical condition (eg, ongoing cancer treatment, severe cardiopulmonary or musculoskeletal disease); magnetic resonance imaging contraindications (eg, pacemakers, aneurysm clips), or stroke or myocardial infarction in the last 6 months that, in the opinion of the primary investigator, would interfere with evaluation of the study or affect the interpretation of the results of the study.

After obtaining informed consent, participants were screened for eligibility by study staff at the Icelandic Heart Association.

Screening for NAFLD

Individuals were screened to assess if they had liver steatosis with a noninvasive ultrasonography-based controlled attenuation parameter (CAP) assessment through a FibroScan device [16]. To avoid overestimation of steatosis, we used 2 probe sizes (medium and extra-large). Individuals who met the full inclusion and exclusion criteria and had a CAP score of >294 dB/m, which represents a high likelihood of >5% liver steatosis, were eligible for participation in the study [17]. Additionally, liver stiffness measurement (LSM) was performed with vibration-controlled transient elastography (VCTE) at screening and at the 12-week follow-up visit. Individuals with an LSM score >9.7 kPa, which represents moderate-to-severe liver fibrosis (grade F3-F4), were referred to a specialist for further evaluation [18]. All individuals with a CAP score >294 dB/m had a magnetic resonance imaging proton density fat fraction (MRI-PDFF) measurement at screening and at the 12-week follow-up visit. MRI-PDFF is considered an emerging biomarker for non-invasive hepatic steatosis assessment as it is accurate, precise, quantitative, and reproducible [19].

The Digital Health Program

The SK-241 digital health program was developed by a multidisciplinary group of experts, including a clinical psychologist, nutritionist, behavioral scientists, medical doctors, and nurses at Sidekick Health. The primary focus of the program was to reduce participants' daily dietary carbohydrate consumption and improve their overall nutrition quality in small, achievable, and sustainable steps (eg, reducing added sugars and processed foods, prioritizing protein, and increasing vegetable consumption). A secondary focus was to increase daily activity levels, improve sleep quality and reduce stress. The user interface with example screenshots from the program is shown in Figure 1.

The program included short daily missions (defined as in-app tasks for the participant to complete) aimed at increasing knowledge about NAFLD and NASH and its contributing factors and improving participants' lifestyles for better metabolic health. The daily missions included watching short educational videos, reading brief informational content, logging meals and beverages by taking a photo of the meal, assessing on a sliding scale how healthy the meal was, and evaluating hunger and satiety before and after the meal. Other missions involved practicing mindfulness and meditation and logging daily energy levels, stress, and sleep quality. The app also provided participants with in-app health coach support (by a live person, not artificial intelligence), which provided weekly feedback on food logs and other in-app activities and opportunities for participants to ask questions as needed. The coaching element was active throughout the 12-week intervention period. Further details of the in-app content and missions are presented in Table 1 and Table S1 in Multimedia Appendix 1.

Figure 1. Example screens of the Sidekick app and the Sidekick-241 NAFLD program user interface. NAFLD: nonalcoholic fatty liver disease.

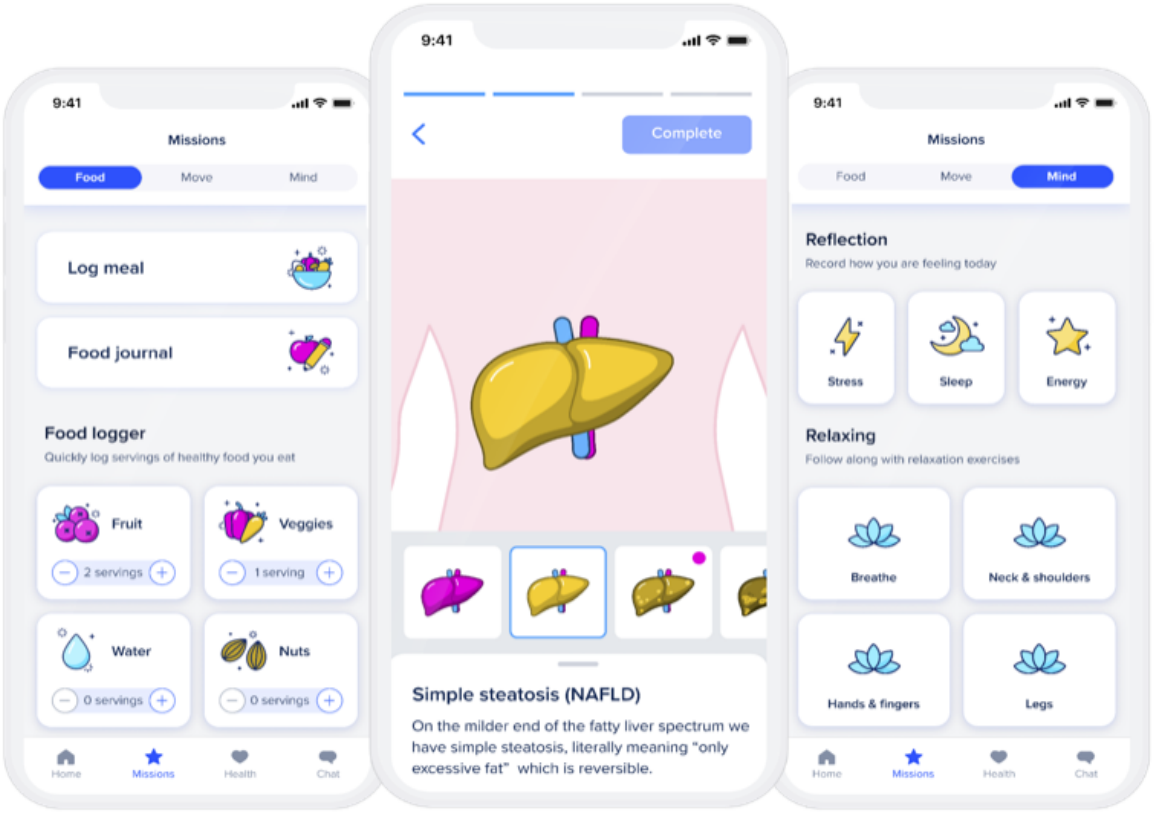


Table 1. The Sidekick-241 program content and descriptions of main missions.

Component	Description
Food journal	Participants were asked to log their meals at least 3 times per week each week. During week 2, individualized goals for gradually reducing carbohydrate intake throughout the program were set based on week 1 consumption.
Step counter	Participants could manually log their steps each day. Individualized goals for increasing steps were set for week 2 based on week 1 step counts.
QoL[a] PROs[b] (stress, sleep, energy)	Participants were prompted to log these measures 2 days per week on a 10-point visual-analog sliding scale.
Surveys	Questions about motivation levels, knowledge and attitudes relating to nutrition and physical activities were administered during weeks 1 and 2 and again during weeks 11 and 12. Questions about current food-related behaviors and potential NAFLD- or NASH-related symptoms were administered every 2 weeks.
Mindfulness	Participants were prompted to complete short mindfulness exercises regularly throughout the program and practice meditation 2 times per week from week 3 onwards.
Coaching	Feedback on weekly in-app activities was provided to participants, particularly on food logs and answers to the in-app surveys. Throughout the program, participants were also able to ask questions as needed and the coach would answer within 24 hours (weekends exempt).

[a]QoL: quality of life.

[b]PROs: patient-reported outcomes.

During the baseline visit, study staff assisted participants with downloading and installing the Sidekick app with the SK-241 program. A short web-based interview with the program's health coach was offered to all participants during the first 2-3 weeks of the study to establish coach connection and accountability and to provide participants with an opportunity to ask questions. During the interview, the primary goals, main concepts, and the program's approach to diet and weight management were explained. In addition, participants' strengths and potential barriers to participation were discussed.

Outcome Measures and Covariates

Primary outcomes were the program's feasibility and acceptability, as assessed by participant retention, engagement, and satisfaction after the 12-week study period. An active participant was defined as one completing at least 1 in-app mission or interacting with the health coach at least once per week. Retention was measured as the number of participants completing the 12-week program, which was defined as being active 9 of 12 weeks. Engagement was measured as the number of participants who were active during the whole 12-week period. Satisfaction with the program was assessed after program completion with the validated mHealth App Usability Questionnaire (MAUQ), which consists of 18 items and has a possible score of 0-7, with 7 being the highest potential score. The scoring can further be divided into 3 subscales reflecting ease of use (5 items), interface and satisfaction (7 items), and usefulness (6 items) [20]. In addition, detailed participant engagement with specific program features was analyzed.

Secondary outcomes were the program's preliminary and potential clinical impact, as measured by weight loss, changes in liver fat, body composition, serum biomarkers, and other cardiometabolic risk factors (eg, blood pressure, waist and hip circumference, and step counts). Participants were assessed at baseline and at a 12-week follow-up visit for demographic information, anthropometric measures, medical history, medications, and adverse events. Liver fat content was measured and quantified at baseline and at 12 weeks using MRI-PDFF with a multiecho chemical shift–encoded gradient-echo sequence [21]. Body composition was assessed at baseline and at 12 weeks with a dual-energy X-ray absorptiometry [22]. Blood pressure was measured using an automatic blood pressure monitor. Blood samples were drawn at baseline and at the 12-week follow-up to measure complete blood count, alanine aminotransferase, aspartate aminotransferase, hemoglobin A_{1c} (HbA_{1c}), fasting glucose and insulin for the homeostatic model assessment of insulin resistance (HOMA-IR), total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, and high-sensitivity C-reactive protein.

Participants were administered the following questionnaires via an electronic patient-reported outcome (PRO) system at baseline and at 12 weeks: the Depression, Anxiety and Stress Scale (DASS-21), the EuroQol-5 Dimension – 5-Level (EQ-5D-5L) index, and the 8-item Morisky Medication Adherence Scale (MMAS-8) [23-25].

For exploratory outcome analysis, study participants were divided into 2 groups depending on how engaged they were with the digital health program. Those using the app 5 or more days per week were defined as highly engaged compared with those using the app less than 5 days per week, and clinical outcomes were compared to assess a potential dose-response relationship.

Statistical Analysis

As this is a feasibility study, a formal sample size calculation was not performed. The researchers aimed for 30-40 participants as this was considered a sufficiently sized sample to obtain information on practical aspects of participants' recruitment, in-app engagement, retention, and rates of acceptance.

Changes in clinical assessments and PROs from baseline to postprogram were calculated as the mean and SD for approximately normally distributed variables (normality was analyzed with the Shapiro-Wilk test) or as the median and IQR for variables that did not satisfy normality criteria. Categorical data were calculated as frequencies and percentages. To compare baseline and postprogram outcomes, paired *t* tests were computed for approximately normally distributed data. In case

the normality assumption was not met, nonparametric tests were computed (Wilcoxon signed-rank tests). Unless otherwise specified, all statistical tests were performed at the 5% (2-sided) significance level. Statistical analysis was performed in Stata (StataCorp) and R (version 4.0.3; R Foundation for Statistical Computing).

All enrolled participants were included in the full analysis set. Missing data were imputed using the last observation carried forward provided that the participant was enrolled in the study and at least one of two measurements (baseline or follow-up) was collected. Moreover, missing baseline measurements in waist circumference, hip circumference, and low-density lipoprotein cholesterol were imputed for 1 participant using the next observation carried backward. The complete case analysis set included participants who attended both the baseline visit and the 12-week follow-up visit.

Ethical Considerations

This study was approved by the National Bioethics Committee of Iceland and the Data Protection Authority (22-075-Vl). All participants provided informed consent before being enrolled in the study. All data was deidentified and analyzed in accordance with institutional protocols. Participants were given the option of seeking reimbursement for travel expenses not exceeding US $150 in total; no other compensation was provided. The study was registered at ClinicalTrials.gov under the trial identifier NCT05426382.

Results

Participant Characteristics

After screening and enrollment, 38 individuals were eligible to participate in the study (Figure 2). The median age of the participants was 59.5 (IQR 46.3-68.8) years, 23 (61%) were women and all were White (Table 2). Of the 38 participants, 17 (45%) had a university degree, none smoked, 34 (90%) had obesity (BMI >30), 19 (50%) had type 2 DM, 27 (71%) had hypertension, 15 (40%) had hypercholesterolemia, and 11 (29%) had a history of cardiovascular disease. Other common comorbidities included hypothyroidism (n=11, 29%), polycystic ovary syndrome (n=4, 11%), and gout (n=2, 5%). In total, 45% (n=17) of participants reported taking antidiabetic medication, 79% (n=30) antihypertensive medication, 37% (n=14) antilipidemic medication, and 37% (n=14) hypothyroid medication. Additionally, 74% (n=28) reported taking other medications, such as proton-pump inhibitors (n=11, 29%), anticoagulants (n=11, 29%), antidepressants (n=8, 21%), vitamin B_{12} (n=5, 13%), nonsteroidal anti-inflammatory medication (n=4, 11%), and antihistamines (n=4, 11%). During the 12-week study period, 5 (13%) participants reported medication changes: 3 (8%) started new medications (one received antibiotics, one received calcium channel blockers, and one vitamin B_{12}) and 2 (5%) reported dosage adjustments (one for diabetes medications and one for beta blockers and antidepressants).

Figure 2. Flowchart of study participants. MRI: magnetic resonance imaging; SK: Sidekick.

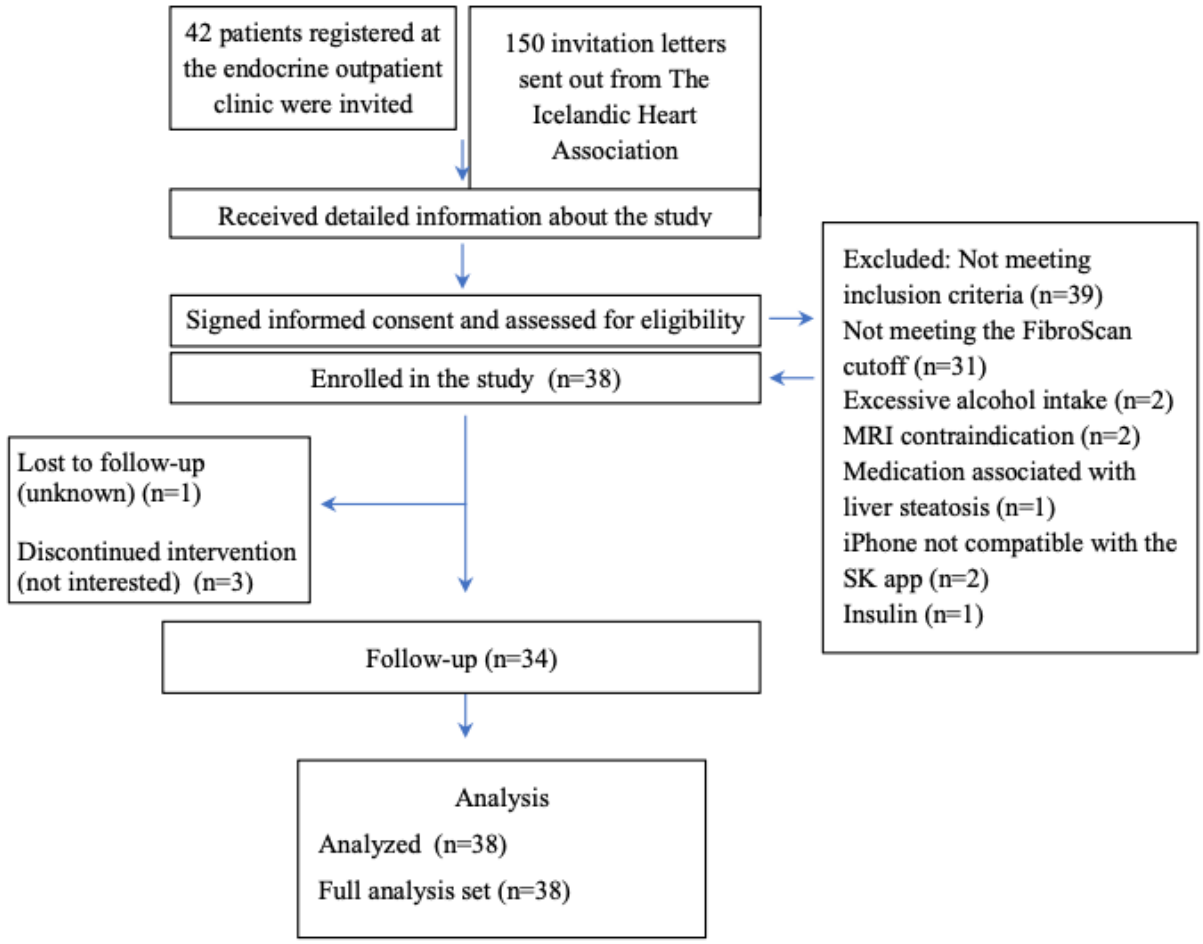


Table 2. Baseline characteristics of the study participants.

Characteristics	Participants (n=38)
Gender, n (%)	
Women	23 (61)
Men	15 (39)
Age (years), median (IQR)	59.5 (46.3-68.8)
Ethnicity: White, n (%)	38 (100)
Work status, n (%)	
Full-time	18 (47)
Part-time	5 (13)
Not in labor market	15 (39)
Pension	11 (29)
Disability	2 (5)
Sick leave	1 (3)
Unemployed	1 (3)
Educational level, n (%)	
University degree	17 (45)
Trades or vocational school or equivalent	12 (32)
Primary education or less	6 (16)
Secondary or matriculate	3 (8)
Smoking status, n (%)	
Current smoker	0 (0)
Never smoked	20 (53)
Former smoker	18 (47)
Comorbidities, n (%)	
Type 2 diabetes	19 (50)
BMI >30	34 (89)
Hypercholesterolemia	15 (39)
Hypertension	27 (71)
Cardiovascular disease	11 (29)
Hypothyroidism	11 (29)
Polycystic ovary disease	4 (11)
Gout	2 (5)
Other	23 (61)

Retention and Engagement in the 12-Week Program

Of the 38 participants, 34 (89%) completed the 12-week program, 29 (76%) were engaged during the whole study period, and 22 (58%) were highly engaged (defined as visiting the app at least 5 days per week) (Table 3). Engagement and retention in the app were similar between those younger or older than 60 years and between men and women (data not shown). Participants were active in-app on a median of 81 (IQR 45.8-84.0) of 84 days or 6.8 (IQR 4.6-7.0) days per week on average and completed an average of 6.9 (SD 2.9) daily missions. Over the course of the study, the health coach sent an average 23.5 (SD 10.3) messages to participants, while participants sent and average of 15.5 (SD 12.4) messages to the coach, who responded within 1.2 (SD 0.9) days. The median MAUQ score was 6.3 (IQR 5.8-6.7) of 7, suggesting high satisfaction with the program among participants.

Table 3. Overall retention, engagement, and satisfaction.

Description	Values
Primary endpoints	
Retention[a], n (%)	34 (89)
Engagement[b], n (%)	29 (76)
Satisfaction[c] median (IQR)	
MAUQ[d] total score	6.3 (5.8-6.7)
Ease of use (mean of MAUQ items 1 to 5)	6.4 (5.6-6.8)
Interface and satisfaction (mean of MAUQ items 6 to 12)	6.3 (5.9-6.9)
Usefulness (mean of MAUQ items 13 to 18)	6.0 (5.5-6.7)
Exploratory engagement metrics	
Average active[e] days per week (0-7), median (IQR)	6.8 (4.6-7.0)
Average total active days (0-84), median (IQR)	81 (45.8-84.0)
Average daily missions completed[f], mean (SD)	6.9 (2.9)
Average daily missions assigned, mean (SD)	5.7 (0.49)
Participants who were active >5 days every week, n (%)	22 (58)
Average number of messages sent by participants, mean (SD)	15.5 (12.4)
Average number of messages received by participants, mean (SD)	23.5 (10.3)

[a]Retention was defined as participants who completed the program, being active for 9 of 12 weeks. Being active was defined as completing at least 1 in-app mission or interacting at least once in that week with the health coach.

[b]Engagement was defined as participants who were active for the entire study period.

[c]Satisfaction was measured using the MAUQ.

[d]MAUQ: mHealth App Usability Questionnaire.

[e]An active day was defined as a day in which the participant completed at least 1 in-app mission or interacted with the coach.

[f]The participants receive daily assigned missions but also had the opportunity to complete additional missions within the app, thereby surpassing the number of assigned missions.

Metabolic Parameters

The mean weight loss was 3.5 (SD 3.7) kg (P<.001), or 3.2% (SD 3.4%) (Table 4). The median body fat percentage changed from 46.6% (IQR 39.4%-52.4%) to 44.3% (IQR 37.8%-52.2%) (P<.001) and the mean fat mass from 50.3 (SD 13.8) kg to 48.1 (SD 14.5) kg (P<.001). These improvements in body composition were accompanied by reduced MRI-PDFF liver fat values: in the full analysis set (n=38), the mean liver fat percentage significantly decreased from 12.3% (SD 7.1%) to 10.1% (SD 6.5%; P<.001), representing a mean relative change of 19.4% (SD 23.9%) (Table 4). In the complete case analysis set (n=34), mean liver fat was reduced from 12.4% (SD 6.9%) to 9.9% (SD 6.3%; P<.001) with a corresponding mean relative change of 21.6% (SD 24.2%).

Table 4. Differences in anthropometric, biochemical, and clinical measurements at baseline and after 12 weeks for the full analysis set (n=38).

	Baseline	Week 12	Change from baseline to week 12	*P* value
Anthropometry				
Weight (kg), mean (SD)	110.0 (18.5)	106.5 (18.4)	3.5 (3.7)	<.001[a]
Relative percentage weight change[b], mean (SD)	N/A[c]	N/A	3.2 (3.4)	N/A
BMI (kg/m^2), mean (SD)	37.6 (5.8)	36.4 (5.8)	1.2 (1.3)	<.001[a]
Waist circumference (cm), mean (SD)	123.8 (12.2)	119.9 (12.2)	4.0 (5.1)	<.001[a]
Hip circumference (cm), mean (SD)	125.1 (14.0)	123.2 (13.3)	1.8 (0.0 to 4.9)	.01[d]
Waist to hip ratio, median (IQR)	1.00 (0.95 to 1.03)	0.99 (0.92 to 1.03)	0.00 (–0.01 to 0.03)	.09[d]
Liver assessment				
Liver fat MRI-PDFF[e] (%), mean (SD)	12.3 (7.1)	10.1 (6.5)	2.2 (2.9)	<.001[a]
Liver fat MRI-PDFF relative change[b] (%), mean (SD)	N/A	N/A	19.4 (23.9)	N/A
Liver stiffness measure (kPa), median (IQR)	6.4 (5.2 to 9.6)	6.6 (5.3 to 8.4)	0.2 (–0.3 to 1.6)	.11[d]
CAP[f] score (dB/m), mean (SD)	343.6 (34.8)	310.3 (47.2)	33.3 (39.7)	<.001[a]
Body composition[g]				
Total body region fat (%), median (IQR)	46.6 (39.4 to 52.4)	44.3 (37.8 to 52.2)	0.9 (1.4)[b]	<.001[a]
Fat mass (kg), mean (SD)	50.3 (13.8)	48.1 (14.5)	2.2 (2.7)	<.001[a]
Lean mass (kg), mean (SD)	56.3 (10.1)	55.6 (9.7)	0.7 (1.7)	.008[a]
Blood pressure (mmHg), mean (SD)				
Systolic	141.4 (17.1)	135.4 (17.3)	6.0 (13.5)	0.009[a]
Diastolic	83.6 (7.4)	82.5 (7.4)	1.2 (7.7)	.36[a]
Biochemical measures				
HbA_{1c}[h] (mmol/mol), median (IQR)	60.0 (56.0 to 66.8)	60.0 (54.3 to 64.0)	0.5 (–0.7 to 3.8)	.03[d]
S-glucose[i] (mmol/L), median (IQR)	6.2 (5.3 to 7.4)	6.3 (5.4 to 6.9)	0.0 (–0.3 to 0.4)	.64[d]
S-insulin[j] (µU/ml), median (IQR)	21.1 (16.4 to 27.9)	19.0 (13.0 to 25.0)	3.2 (0.0 to 5.4)	.003[d]
HOMA-IR[k] (mmol/L), median (IQR)	5.8 (4.3 to 8.4)	4.8 (3.6 to 7.2)	0.4 (–0.2 to 2.1)	.02[d]
Total cholesterol (mmol/L), mean (SD)	4.9 (1.3)	4.8 (1.2)	0.0 (–0.2 to 0.2)	>.99[d]
LDL-C[l] (mmol/L), mean (SD)	2.9 (1.1)	2.9 (1.1)	–0.1 (–0.3 to 0.1)	.18[d]
HDL-C[m] (mmol/L), mean (SD)	1.11 (0.23)	1.12 (0.19)	–0.01 (0.12)	.56[a]
Triglycerides (mmol/L), median (IQR)	1.88 (1.35 to 2.45)	1.68 (1.21 to 1.90)	0.14 (0.00 to 0.47)	.003[d]
hs-CRP[n] (mg/L), median (IQR)	3.0 (1.2 to 5.2)	2.5 (1.1 to 3.9)	0.1 (–0.1 to 0.7)	.14[d]
ALAT[o] (IU/L), median (IQR)	21.4 (18.2 to 30.2)	23.2 (18.4 to 32.0)	0.0 (–6.8 to 2.8)	.37[d]
ASAT[p] (IU/L), median (IQR)	20.8 (17.9 to 24.8)	22.3 (18.0 to 25.5)	0.4 (–2.5 to 2.5)	.53[d]
FIB-4[q] Index, median (IQR)	1.08 (0.78 to 1.34)	1.08 (0.75 to 1.21)	0.01 (–0.06 to 0.07)	.58[d]

[a]Analyzed with a paired *t* test.

[b]Percentage change calculated as the average over individual relative changes.

[c]N/A: not applicable.

[d]Analyzed with a Wilcoxon signed-rank test.

[e]MRI-PDFF: magnetic resonance imaging proton density fat fraction.

[f]CAP: controlled attenuation parameter.

[g]Measured by dual-energy ray absorptiometry.

[h]HbA_{1c}: glycated hemoglobin A_{1c}.

[i]s-glucose: Serum glucose.

[j]s-insulin: Serum insulin.

[k]HOMA-IR: homeostatic model assessment of insulin resistance.

[l]LDL-C: low-density lipoprotein cholesterol.

[m]HDL-C: high-density lipoprotein cholesterol.

[n]hs-CRP high-sensitivity C-reactive protein.

[o]ALAT: alanine aminotransferase.

[p]ASAT: aspartate aminotransferase.

[q]FIB-4: index for liver fibrosis.

During the study, the distribution of steatosis levels changed. At baseline, the 10%-15% liver steatosis category had the highest frequency with 32% (n=12) of participants. At follow-up, the 5%-10% liver steatosis category had the highest frequency with 34% (n=13) of participants (Table 5). We additionally found a significant correlation between weight loss and absolute (r=0.48, P=.004) and relative (r=0.72, P<.001) liver fat changes measured by MRI-PDFF.

According to the FIB-4 data, 4 of the 38 participants were classified as having a high risk of fibrosis at baseline, of which 1 individual regressed to intermediate risk at the 12-week follow-up visit. Most participants (n=27, 71%) had a low risk of fibrosis at baseline according to the FIB-4. At the 12-week follow-up, this percentage had gone up to 79% (n=30).

Mean systolic blood pressure significantly decreased by 6.0 (SD 13.5) mmHg (P=.009), and this was not explained by changes in medication or medication adherence (Table 4). There was no significant difference in diastolic blood pressure.

Participants recorded on average 3085 (SD 2246) daily steps in the first week and 4664 (SD 3780) daily steps in the last week, representing a significant increase of 1579 steps per day (P=.02).

While participants' average baseline fasting insulin and HOMA-IR levels indicated insulin resistance, we found a significant decrease in serum insulin levels (median 3.2, IQR 0.0-5.4 µU/ml; P=.003), HOMA-IR levels (median 0.4, IQR –0.2 to 2.1 mmol/L; P=.02), and HbA_{1c} levels (median 0.5, IQR –0.7 to 3.8 mmol/mol; P=.03) (Table 4), suggesting improved glycemic control. In addition, triglyceride levels significantly decreased by a median of 0.14 (IQR 0.00-0.47) mmol/L (P=.003), and median high-sensitivity C-reactive protein levels decreased from 3.0 (IQR 1.2-5.2) mg/L to 2.5 (IQR 1.1-3.9) mg/L (P=.14) (Table 4), representing improvements in those cardiovascular risk factors.

We did not find any significant change in cholesterol levels, nor any significant changes in PRO scores of health-related quality of life, mental health, or medication adherence from preprogram to postprogram (Table 6).

Table 5. Distributions of liver fat percentage categories based on magnetic resonance imaging proton density fat fraction liver fat values at baseline and at the 12-week follow-up (n=38). All participants with >5% liver fat at baseline had stage 1 steatosis according to the standardized Nonalcoholic Steatohepatitis Clinical Research Network histologic scoring system for nonalcoholic fatty liver disease [26].

Liver fat category (%)[a]	Participants at baseline, n	Participants at week 12, n
<5	6	9
5-10	9	13
10-15	12	7
15-20	5	5
20-25	4	3
25-30	2	1

[a]Limits for the presented ranges correspond to values greater than or equal to for lower limits and less than for upper limits.

Table 6. Differences in patient reported outcomes (PROs) at baseline and after 12 weeks for the full analysis set (n=38).

PROs	Baseline	Week 12	Change from baseline to week 12	*P* value[a]
EQ-5D-5L index[b]	0.8 (0.7 to 0.9)	0.9 (0.8 to 1.0)	0.0 (–0.1 to 0.0)	.36
DASS-21[c], median (IQR)				
Total score (0 to 56)	5.5 (2.0 to 13.0)	6.0 (2.0 to 11.0)	0.0 (–2.0 to 2.0)	.95
Depression score	1.0 (0.3 to 4.5)	1.0 (0.0 to 3.0)	0.0 (0.0 to 1.0)	—[d]
Anxiety score	1.0 (0.0 to 2.0)	1.0 (0.0 to 2.0)	0.0 (–0.7 to 1.0)	—
Stress score	3.0 (1.0 to 6.0)	3.0 (0.3 to 6.0)	0.0 (–1.0 to 1.0)	—
MMAS-8[e]				
Total score (0 to 8)	7.0 (6.8 to 8.0)	7.0 (6.8 to 8.0)	0.0 (–0.7 to 0.0)	.68
High adherence (=8), n (%)	14 (37)	15 (39)	N/A[f]	—
Moderate adherence (6 to 7), n (%)	15 (39)	15 (39)	N/A	—
Low adherence (<6), n (%)	9 (24)	8 (21)	N/A	—

[a]Analyzed with Wilcoxon signed-rank test.

[b]EQ-5D-5L index: EuroQol-5 Dimension – 5-Level index.

[c]DASS-21: Depression, Anxiety, and Stress Scale - 21 Items.

[d]Not available.

[e]MMAS-8: 8-item Morisky Medication Adherence Scale.

[f]N/A: not applicable.

Associations Between App Engagement and Clinical Outcomes

An exploratory analysis was performed to assess the relationship between participants' in-app activity and their clinical outcomes. We found that participants who were highly engaged (visited the app at least 5 days per week) had greater weight loss and liver fat reduction (Table S2 in Multimedia Appendix 1) compared with those who were less engaged. In a complete case analysis, participants who were highly engaged (n=22) lost on average 5.1 (SD 3.8) kg and achieved a 27.5% relative reduction in liver fat, while those who were active on fewer than 5 days a week (n=12) lost on average 1.8 (SD 2.2) kg and achieved 10.8% relative reduction in liver fat. Moreover, highly engaged participants were significantly more likely to achieve a relative weight loss of at least 3% (*P*=.001) or 5% (*P*=.02) compared with those who were less engaged (Fisher exact tests). Taken together, these results suggest that higher engagement with the digital program may be associated with improved metabolic health.

Adverse Events

In total, 9 adverse events were reported, all of which were of mild to moderate intensity with no serious adverse events (Table S3 in Multimedia Appendix 1). No adverse events were considered to have a causal relationship to the digital health program, as assessed by the primary investigator.

Discussion

Principal Findings

This study demonstrated that the 12-week-long digital health program, SK-241, was feasible given its high retention, engagement, and satisfaction among people with NAFLD. Cardiometabolic health and liver-specific outcomes improved over the 12-week study period with a significant weight loss and reductions in fat mass, liver fat, systolic blood pressure, triglycerides, insulin, and HbA_{1c} levels.

Digital behavioral programs can be effective at targeting weight loss among people with chronic conditions [27]. Increasing evidence shows that programs—whether digital or face-to-face—with a holistic approach can also be effective for people with NAFLD, where weight loss is a major component of disease management. A recent randomized controlled study from Singapore including 108 adults with NAFLD randomized either to lifestyle advice by a trained nurse or using a lifestyle mobile app in addition to receiving advice by a dietitian showed that the mobile app group had a 5-fold higher likelihood of achieving ≥5% weight loss compared with the control group at 6 months [28].

Previous studies suggest that digital solutions can be as effective as face-to-face behavioral change programs, but engagement with the digital program is an important component of efficacy [29]. Indeed, an important finding of this study was the correlation between participants' in-app engagement and their clinical outcomes; this shows that maintaining engagement and interest is key to reaching the desired clinical improvements. Program engagement may be influenced by several factors, such as recruitment methods, participant characteristics, app design, and the level of support, such as coaching [30,31]. Coaching in particular may be essential to drive engagement as it encourages accountability and may increase motivation [30]. The regular contact that participants had with the coach in the SK-241 program may have contributed to the low attrition and high engagement in this study. Should larger implementation of this

intervention take place, then coaching would be an integral part, at least in the initial stages of the program.

Lifestyle interventions consisting of diet, exercise, and weight loss are recommended to individuals with NAFLD according to treatment guidelines [2]. The primary driver of NAFLD is overnutrition, which causes expansion of adipose deposits and macrophage infiltration into the visceral adipose tissue, creating a proinflammatory state that promotes insulin resistance [32,33]. The resulting imbalance in lipid metabolism leads to the formation of lipotoxic lipids that contribute to cellular stress, including oxidative stress, inflammasome activation, and apoptotic cell death [34,35]. Central obesity is also an important driver of insulin resistance and proinflammatory signaling [36]. In this 12-week study, the mean waist circumference was significantly reduced by 4.0 cm, and body weight by 3.2% on average; these are encouraging results, considering that a 3%-5% weight loss can lead to a reduction in hepatic steatosis [37]. In addition, we found a correlation between weight loss and MRI-PDFF liver fat fraction changes. This is in line with a previous report of greater weight loss leading to more significant improvements in liver histopathology, and studies have shown that a ≥30% relative decline in liver fat by MRI-PDFF is associated with histopathological improvements in NASH [38-40]. Participants in this study were able to decrease their waist circumference and body weight and had an average of around a 20% relative reduction in liver fat by MRI-PDFF, with a subset of participants achieving a 30% relative decline, which might lead to improved NAFLD and NASH histopathology.

Despite the risk of progressive liver disease, the leading cause of death in people with NAFLD is cardiovascular disease [10]. This is likely due to risk factors that are shared between NAFLD and cardiovascular diseases, although it is unclear to what extent NAFLD has a direct causative role in the development of cardiovascular disease [41]. Therefore, it was important to see significant improvement in cardiovascular risk factors in our study, such as a decrease in systolic blood pressure, triglycerides, insulin, and HbA_{1c}. The increased physical activity in our study as measured by the in-app step counter and the correlation between in-app activity and weight loss suggest that the digital program may successfully engage participants in behaviors that lead to more weight loss, which in turn may hypothetically improve liver function and glycemic control. Regular tracking of meals and physical activity and completing the in-app PROs may help people become more aware of their habits, while the education and the coach's feedback and support may give them the necessary tools to change their behaviors. We did not find any significant changes in the PRO scores of health-related quality of life and mental health, which was most likely due to the short duration of the study and the small size of the cohort.

Furthermore, studies have shown that health care utilization and expenditure are particularly high among people with NAFLD and NASH [42,43]. Therefore, there is a great need for early identification and effective management of people with NAFLD to minimize the comorbidity burden and health care costs.

The fibrosis risk among study participants was assessed both with a VCTE FibroScan LSM and by calculating the FIB-4 index score from participants' age and the serum alanine aminotransferase, aspartate aminotransferase, and platelet count. The results indicated that a few participants had an intermediate to high risk of having liver fibrosis (data not shown) and could be referred to as probable patients with NASH, thereby suggesting that the digital health program might be feasible for individuals with NASH in addition to those with NAFLD. However, both of these measurements have their limitations and need to be interpreted cautiously. VCTE can rule out advanced fibrosis but often leads to false positive results in NAFLD, while the FIB-4 score might overestimate fibrosis in populations older than 65 years and is considered to have a low positive predictive value for identifying advanced fibrosis [44,45].

Strengths and Limitations

A strength of this study was the high engagement and completion rate, as these are well known issues of digital health programs [46]. In addition, the holistic nature of the program, developed by a multidisciplinary team of experts and focused on multiple aspects of participants' lifestyle, combined with the regular support provided by the coach, can be considered a strength. A further strength was the length of the program, which allowed sufficient time to assess meaningful changes in engagement and clinical outcomes.

Limitations of this study included the single-arm design, which limits the interpretation and generalizability of our findings. The lack of a control group made it difficult to directly infer the clinical benefit of digital program, thus the secondary outcomes relating to clinical efficacy should be interpreted with caution. The observed clinical improvements should also be interpreted in context with the short duration of the health program, as sustaining improvements can be challenging after short-term behavioral interventions. It should also be acknowledged that a seasonal increase in activity levels may have contributed to the observed changes, as the study began in early summer when people tend to be more physically active. Furthermore, all the participants were White and around 50% had a relatively high education level. Higher education level has been associated with a lower burden of traditional cardiovascular risk factors [47]. Previous studies have shown an association between socioeconomic status and NAFLD, where poverty seems to be a risk factor for developing NAFLD independent of other known risk factors, such as type 2 DM and obesity, and food insecurity is associated with developing NAFLD and advanced fibrosis [48]. Education and smoking status may have affected engagement with the digital health program and, therefore, the generalizability of these results to a wider population may be limited and future trials should recruit a more diverse group of participants to assess the efficacy of the program [49].

Conclusions

The 12-week-long digital health program was feasible for individuals with NAFLD, showing high user engagement, retention, and satisfaction. Improved liver-specific and cardiometabolic health was observed and more engaged

participants showed greater improvements. This NAFLD digital health program could provide a new tool to improve health outcomes in people with NAFLD.

Acknowledgments

Judit Mészáros and Joanna McCarter provided medical writing and editing support. We would also like to thank the staff at the Icelandic Heart Association who helped with the study procedures, as well as the people who participated in the study. The study was funded by Sidekick Health. Use of the MMAS-8 is protected by US and international copyright laws. Permission for use is required. A Licensure agreement is available from: MMAR, LLC., www.moriskyscale.com.

Data Availability

The data sets generated or analyzed during this study are not publicly available due to restrictions in the informed consent form. Additional summary statistics will be provided upon on reasonable request.

Authors' Contributions

SB, HU, EFG, KS, TG, TK, SO, and VG contributed to conceptualization. SB, HU, EFG, KS, AI, BD, GEAM, TG, GB, SS, SO, and VG contributed to the methodology. SB, HU, EFG, AI, BD, SS, SO, and VG contributed to the investigation. SB and HU wrote the original draft. SB, HU, EFG, KS, AI, BD, GEAM, TG, TK, GB, SS, SO, and VG revised and edited the manuscript. SO acquired funding. SO and VG procured resources and supervised the study.

Conflicts of Interest

HU, EFG, AI, BD, KS, TK, GEAM, and TG are employed by Sidekick Health. SO is an employee and cofounder of Sidekick Health. SB received consultancy fees from Sidekick Health during the study period. SS, GB, and VG have no competing interests to declare.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 17 KB - cardio_v8i1e52576_app1.docx]

References

1. Rinella ME. Nonalcoholic fatty liver disease: a systematic review. JAMA 2015 Jun 09;313(22):2263-2273. [doi: 10.1001/jama.2015.5370] [Medline: 26057287]
2. European Association for the Study of the Liver (EASL), European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016 Jun;64(6):1388-1402. [doi: 10.1016/j.jhep.2015.11.004] [Medline: 27062661]
3. Budd J, Cusi K. Nonalcoholic fatty liver disease: what does the primary care physician need to know? Am J Med 2020 May;133(5):536-543. [doi: 10.1016/j.amjmed.2020.01.007] [Medline: 32017891]
4. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016 Jul;64(1):73-84. [doi: 10.1002/hep.28431] [Medline: 26707365]
5. Younossi ZM, Golabi P, de Avila L, Paik JM, Srishord M, Fukui N, et al. The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: a systematic review and meta-analysis. J Hepatol 2019 Oct;71(4):793-801. [doi: 10.1016/j.jhep.2019.06.021] [Medline: 31279902]
6. Polyzos SA, Kountouras J, Mantzoros CS. Obesity and nonalcoholic fatty liver disease: from pathophysiology to therapeutics. Metabolism 2019 Mar;92:82-97. [doi: 10.1016/j.metabol.2018.11.014] [Medline: 30502373]
7. Fernando DH, Forbes JM, Angus PW, Herath CB. Development and progression of non-alcoholic fatty liver disease: the role of advanced glycation end products. Int J Mol Sci 2019 Oct 11;20(20):5037 [FREE Full text] [doi: 10.3390/ijms20205037] [Medline: 31614491]
8. Galiero R, Caturano A, Vetrano E, Cesaro A, Rinaldi L, Salvatore T, et al. Pathophysiological mechanisms and clinical evidence of relationship between nonalcoholic fatty liver disease (NAFLD) and cardiovascular disease. Rev Cardiovasc Med 2021 Sep 24;22(3):755-768 [FREE Full text] [doi: 10.31083/j.rcm2203082] [Medline: 34565074]
9. Byrne CD, Targher G. NAFLD as a driver of chronic kidney disease. J Hepatol 2020 Apr;72(4):785-801. [doi: 10.1016/j.jhep.2020.01.013] [Medline: 32059982]
10. Younossi ZM. Long-term outcomes of nonalcoholic fatty liver disease: from nonalcoholic steatohepatitis to nonalcoholic steatofibrosis. Clin Gastroenterol Hepatol 2017 Aug;15(8):1144-1147. [doi: 10.1016/j.cgh.2017.05.029] [Medline: 28549675]

11. Godoy-Matos AF, Silva Júnior WS, Valerio CM. NAFLD as a continuum: from obesity to metabolic syndrome and diabetes. Diabetol Metab Syndr 2020;12:60 [FREE Full text] [doi: 10.1186/s13098-020-00570-y] [Medline: 32684985]
12. Cusi K, Isaacs S, Barb D, Basu R, Caprio S, Garvey WT, et al. American Association of Clinical Endocrinology clinical practice guideline for the diagnosis and management of nonalcoholic fatty liver disease in primary care and endocrinology clinical settings: co-sponsored by the American Association for the Study of Liver Diseases (AASLD). Endocr Pract 2022 May;28(5):528-562. [doi: 10.1016/j.eprac.2022.03.010] [Medline: 35569886]
13. Mantovani A, Dalbeni A. Treatments for NAFLD: state of art. Int J Mol Sci 2021 Feb 26;22(5):2350 [FREE Full text] [doi: 10.3390/ijms22052350] [Medline: 33652942]
14. Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. Obesity Reviews 2011 Aug 05;12(11):912-934. [doi: 10.1111/j.1467-789x.2011.00915.x]
15. Svansdottir E, Denollet J, Thorsson B, Gudnason T, Halldorsdottir S, Gudnason V, et al. Association of type D personality with unhealthy lifestyle, and estimated risk of coronary events in the general Icelandic population. Eur J Prev Cardiol 2013 Apr;20(2):322-330. [doi: 10.1177/2047487312441723] [Medline: 22383854]
16. Zhang X, Wong GL, Wong VW. Application of transient elastography in nonalcoholic fatty liver disease. Clin Mol Hepatol 2020 Apr;26(2):128-141 [FREE Full text] [doi: 10.3350/cmh.2019.0001n] [Medline: 31696690]
17. Ferraioli G. Quantitative assessment of liver steatosis using ultrasound controlled attenuation parameter (Echosens). J Med Ultrason (2001) 2021 Oct;48(4):489-495 [FREE Full text] [doi: 10.1007/s10396-021-01106-1] [Medline: 34132934]
18. Eddowes PJ, Sasso M, Allison M, Tsochatzis E, Anstee QM, Sheridan D, et al. Accuracy of FibroScan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with nonalcoholic fatty liver disease. Gastroenterology 2019 May;156(6):1717-1730. [doi: 10.1053/j.gastro.2019.01.042] [Medline: 30689971]
19. Caussy C, Reeder SB, Sirlin CB, Loomba R. Noninvasive, quantitative assessment of liver fat by MRI-PDFF as an endpoint in NASH trials. Hepatology 2018 Aug;68(2):763-772 [FREE Full text] [doi: 10.1002/hep.29797] [Medline: 29356032]
20. Zhou L, Bao J, Setiawan IMA, 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]
21. França M, Alberich-Bayarri Á, Martí-Bonmatí L, Oliveira P, Costa FE, Porto G, et al. Accurate simultaneous quantification of liver steatosis and iron overload in diffuse liver diseases with MRI. Abdom Radiol (NY) 2017 May;42(5):1434-1443. [doi: 10.1007/s00261-017-1048-0] [Medline: 28110367]
22. Hull H, He Q, Thornton J, Javed F, Allen L, Wang J, et al. iDXA, Prodigy, and DPXL dual-energy X-ray absorptiometry whole-body scans: a cross-calibration study. J Clin Densitom 2009;12(1):95-102 [FREE Full text] [doi: 10.1016/j.jocd.2008.09.004] [Medline: 19028125]
23. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008 May 02;10(5):348-354 Retracted in: J Clin Hypertens (Greenwich). 2023 Sep;25(9):889. J Clin Hypertens (Greenwich). 2023 Sep;25(9):890. [FREE Full text] [doi: 10.1111/j.1751-7176.2008.07572.x] [Medline: 18453793]
24. Berlowitz DR, Foy CG, Kazis LE, Bolin LP, Conroy MB, Fitzpatrick P, for the SPRINT Research Group. Effect of intensive blood-pressure treatment on patient-reported outcomes. N Engl J Med 2017 Aug 24;377(8):733-744 [FREE Full text] [doi: 10.1056/NEJMoa1611179] [Medline: 28834483]
25. Bress AP, Bellows BK, King JB, Hess R, Beddhu S, Zhang Z, for the SPRINT Research Group. Cost-effectiveness of intensive versus standard blood-pressure control. N Engl J Med 2017 Aug 24;377(8):745-755 [FREE Full text] [doi: 10.1056/NEJMsa1616035] [Medline: 28834469]
26. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005 Jun;41(6):1313-1321. [doi: 10.1002/hep.20701] [Medline: 15915461]
27. Fakih El Khoury C, Karavetian M, Halfens RJ, Crutzen R, Khoja L, Schols JM. The effects of dietary mobile apps on nutritional outcomes in adults with chronic diseases: a systematic review and meta-analysis. J Acad Nutr Diet 2019 Apr;119(4):626-651. [doi: 10.1016/j.jand.2018.11.010] [Medline: 30686742]
28. Lim SL, Johal J, Ong KW, Han CY, Chan YH, Lee YM, et al. Lifestyle intervention enabled by mobile technology on weight loss in patients with nonalcoholic fatty liver disease: randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 13;8(4):e14802 [FREE Full text] [doi: 10.2196/14802] [Medline: 32281943]
29. Mazzotti A, Caletti MT, Brodosi L, Di Domizio S, Forchielli ML, Petta S, et al. An internet-based approach for lifestyle changes in patients with NAFLD: two-year effects on weight loss and surrogate markers. J Hepatol 2018 Nov;69(5):1155-1163 [FREE Full text] [doi: 10.1016/j.jhep.2018.07.013] [Medline: 30290973]
30. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
31. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]

32. Kneeman JM, Misdraji J, Corey KE. Secondary causes of nonalcoholic fatty liver disease. Therap Adv Gastroenterol 2011 Dec 02;5(3):199-207. [doi: 10.1177/1756283x11430859]
33. Zatterale F, Longo M, Naderi J, Raciti GA, Desiderio A, Miele C, et al. Chronic adipose tissue inflammation linking obesity to insulin resistance and type 2 diabetes. Front Physiol 2019;10:1607 [FREE Full text] [doi: 10.3389/fphys.2019.01607] [Medline: 32063863]
34. Sanyal AJ. Past, present and future perspectives in nonalcoholic fatty liver disease. Nat Rev Gastroenterol Hepatol 2019 Jun;16(6):377-386. [doi: 10.1038/s41575-019-0144-8] [Medline: 31024089]
35. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. Nat Med 2018 Jul;24(7):908-922 [FREE Full text] [doi: 10.1038/s41591-018-0104-9] [Medline: 29967350]
36. Hardy OT, Czech MP, Corvera S. What causes the insulin resistance underlying obesity? Curr Opin Endocrinol Diabetes Obes. Curr Opin Endocrinol Diabetes Obes 2012;19(2):81-87. [doi: 10.1097/med.0b013e3283514e13]
37. Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. Hepatology 2018 Jan;67(1):328-357. [doi: 10.1002/hep.29367] [Medline: 28714183]
38. Vilar-Gomez E, Martinez-Perez Y, Calzadilla-Bertot L, Torres-Gonzalez A, Gra-Oramas B, Gonzalez-Fabian L, et al. Weight loss through lifestyle modification significantly reduces features of nonalcoholic steatohepatitis. Gastroenterology 2015 Aug;149(2):367-378. [doi: 10.1053/j.gastro.2015.04.005] [Medline: 25865049]
39. Loomba R, Neuschwander-Tetri BA, Sanyal A, Chalasani N, Diehl AM, Terrault N, et al. Multicenter validation of association between decline in MRI-PDFF and histologic response in NASH. Hepatology 2020 Oct;72(4):1219-1229 [FREE Full text] [doi: 10.1002/hep.31121] [Medline: 31965579]
40. Stine JG, Munaganuru N, Barnard A, Wang JL, Kaulback K, Argo CK, et al. Change in MRI-PDFF and histologic response in patients with nonalcoholic steatohepatitis: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2021 Nov;19(11):2274-2283.e5 [FREE Full text] [doi: 10.1016/j.cgh.2020.08.061] [Medline: 32882428]
41. Targher G, Byrne CD, Tilg H. NAFLD and increased risk of cardiovascular disease: clinical associations, pathophysiological mechanisms and pharmacological implications. Gut 2020 Sep;69(9):1691-1705. [doi: 10.1136/gutjnl-2020-320622] [Medline: 32321858]
42. Younossi ZM, Tampi R, Priyadarshini M, Nader F, Younossi IM, Racila A. Burden of illness and economic model for patients with nonalcoholic steatohepatitis in the United States. Hepatology 2019 Feb;69(2):564-572. [doi: 10.1002/hep.30254] [Medline: 30180285]
43. Younossi ZM, Blissett D, Blissett R, Henry L, Stepanova M, Younossi Y, et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. Hepatology 2016 Nov;64(5):1577-1586. [doi: 10.1002/hep.28785] [Medline: 27543837]
44. European Association for the Study of the Liver, Clinical Practice Guideline Panel. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. J Hepatol 2021 Sep;75(3):659-689. [doi: 10.1016/j.jhep.2021.05.025] [Medline: 34166721]
45. Sumida Y, Yoneda M, Tokushige K, Kawanaka M, Fujii H, Yoneda M, et al. FIB-4 first in the diagnostic algorithm of metabolic-dysfunction-associated fatty liver disease in the era of the global metabodemic. Life (Basel) 2021 Feb 14;11(2):143 [FREE Full text] [doi: 10.3390/life11020143] [Medline: 33672864]
46. 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 [FREE Full text] [doi: 10.2196/35120] [Medline: 35471414]
47. Andersen K, Aspelund T, Gudmundsson EF, Sigurdsson G, Sigurdsson S, Björnsdottir G, et al. The relationship between lack of educational attainment, cardiovascular risk factors, atherosclerosis and coronary artery disease. Laeknabladid 2022 Jul;108:346-355. [doi: 10.17992/lbl.2022.0708.701] [Medline: 35943050]
48. Frohme J, Tacke F. The socioeconomic aspects of nonalcoholic fatty liver disease: food insecurity as a novel risk factor for steatosis and liver fibrosis. Hepatobiliary Surg Nutr 2020 Aug;9(4):543-545 [FREE Full text] [doi: 10.21037/hbsn.2020.03.03] [Medline: 32832515]
49. Jiandani D, Wharton S, Rotondi MA, Ardern CI, Kuk JL. Predictors of early attrition and successful weight loss in patients attending an obesity management program. BMC Obes 2016;3:14 [FREE Full text] [doi: 10.1186/s40608-016-0098-0] [Medline: 26966544]

Abbreviations

CAP: controlled attenuation parameter
DASS-21: Depression, Anxiety, Stress Scale, 21 Items
DM: diabetes mellitus
EQ-5D-5L index: EuroQol-5 Dimension – 5-Level index
HbA1c: hemoglobin A1c
HOMA-IR: homeostatic model assessment for insulin resistance

LSM: liver stiffness measurement
MAUQ: mHealth App Usability Questionnaire
MMAS-8: 8-item Morisky Medication Adherence Scale
MRI-PDFF: magnetic resonance imaging proton density fat fraction
NAFLD: nonalcoholic fatty liver disease
NASH: nonalcoholic steatohepatitis
ppt: percentage points
PRO: patient-reported outcome
s-insulin: Serum insulin
Total-C: total cholesterol
VCTE: vibration-controlled transient elastography

Edited by A Mavragani; submitted 12.09.23; peer-reviewed by G Marchesini, J Shull; comments to author 30.11.23; revised version received 15.12.23; accepted 26.12.23; published 15.02.24.

Please cite as:
Björnsdottir S, Ulfsdottir H, Gudmundsson EF, Sveinsdottir K, Isberg AP, Dobies B, Akerlie Magnusdottir GE, Gunnarsdottir T, Karlsdottir T, Bjornsdottir G, Sigurdsson S, Oddsson S, Gudnason V
User Engagement, Acceptability, and Clinical Markers in a Digital Health Program for Nonalcoholic Fatty Liver Disease: Prospective, Single-Arm Feasibility Study
JMIR Cardio 2024;8:e52576
URL: https://cardio.jmir.org/2024/1/e52576
doi:10.2196/52576
PMID:38152892

Original Paper

Cognitive Behavioral Therapy for Symptom Preoccupation Among Patients With Premature Ventricular Contractions: Nonrandomized Pretest-Posttest Study

Björn E Liliequist[1], MSc; Josefin Särnholm[1], PhD; Helga Skúladóttir[2,3], MD; Eva Ólafsdóttir[2,3], RN; Brjánn Ljótsson[1], PhD; Frieder Braunschweig[2,3], MD, PhD

[1]Department of Clinical Neuroscience, Division of Psychology, Karolinska Institutet, Stockholm, Sweden

[2]Department of Medicine, Cardiology Unit, Karolinska Institutet, Stockholm, Sweden

[3]Department of Cardiology, Karolinska University Hospital, Stockholm, Sweden

Corresponding Author:
Björn E Liliequist, MSc
Department of Clinical Neuroscience
Division of Psychology
Karolinska Institutet
Nobels väg 9
Stockholm, 171 65
Sweden
Phone: 46 08 524 800 00
Fax: 46 08 31 11 01
Email: bjorn.liliequist@ki.se

Abstract

Background: Premature ventricular contractions (PVCs) are a common cardiac condition often associated with disabling symptoms and impaired quality of life (QoL). Current treatment strategies have limited effectiveness in reducing symptoms and restoring QoL for patients with PVCs. Symptom preoccupation, involving cardiac-related fear, hypervigilance, and avoidance behavior, is associated with disability in other cardiac conditions and can be effectively targeted by cognitive behavioral therapy (CBT).

Objective: The aim of this study was to evaluate the effect of a PVC-specific CBT protocol targeting symptom preoccupation in patients with symptomatic idiopathic PVCs.

Methods: Nineteen patients diagnosed with symptomatic idiopathic PVCs and symptom preoccupation underwent PVC-specific CBT over 10 weeks. The treatment was delivered by a licensed psychologist via videoconference in conjunction with online text-based information and homework assignments. The main components of the treatment were exposure to cardiac-related symptoms and reducing cardiac-related avoidance and control behavior. Self-rated measures were collected at baseline, post treatment, and at 3- and 6-month follow-ups. The primary outcome was PVC-specific QoL at posttreatment assessment measured with a PVC-adapted version of the Atrial Fibrillation Effects on Quality of Life questionnaire. Secondary measures included symptom preoccupation measured with the Cardiac Anxiety Questionnaire. PVC burden was evaluated with 5-day continuous electrocardiogram recordings at baseline, post treatment, and 6-month follow-up.

Results: We observed large improvements in PVC-specific QoL (Cohen d=1.62, P<.001) and symptom preoccupation (Cohen d=1.73, P<.001) post treatment. These results were sustained at the 3- and 6-month follow-ups. PVC burden, as measured with 5-day continuous electrocardiogram, remained unchanged throughout follow-up. However, self-reported PVC symptoms were significantly lower at posttreatment assessment and at both the 3- and 6-month follow-ups. Reduction in symptom preoccupation had a statistically significant mediating effect of the intervention on PVC-specific QoL in an explorative mediation analysis.

Conclusions: This uncontrolled pilot study shows preliminary promising results for PVC-specific CBT as a potentially effective treatment approach for patients with symptomatic idiopathic PVCs and symptom preoccupation. The substantial improvements in PVC-specific QoL and symptom preoccupation, along with the decreased self-reported PVC-related symptoms warrant further investigation in a larger randomized controlled trial.

Trial Registration: ClinicalTrials.gov NCT05087238; https://clinicaltrials.gov/study/NCT05087238

(JMIR Cardio 2024;8:e53815) doi:10.2196/53815

KEYWORDS

premature ventricular contractions; quality of life; symptom preoccupation; cognitive behavioral therapy: CBT

Introduction

Premature ventricular contractions (PVCs) are a common type of cardiac arrhythmia with a prevalence of 69%-99.5% in the adult population [1,2]. PVCs are commonly asymptomatic but can result in mild to disabling symptoms due to palpitations, dyspnea, presyncope, and fatigue [3]. In the absence of structural heart disease or inherited ion channelopathies, PVCs are referred to as idiopathic and considered benign, although a high PVC burden may induce a cardiomyopathy in some patients [3,4]. Medical treatment with beta-blockers or nondihydropyridine calcium channel blockers can decrease the PVC burden and provide symptomatic improvement; however, this treatment is ineffective in a large portion of patients and may cause side effects. Catheter ablation is the most effective approach to abolish PVCs and has been emphasized in recent guidelines as the recommended first-line treatment for symptomatic idiopathic PVCs [5]. However, ablation may not be readily available or may be offered with some reservation owing to the risk of rare but potentially life-threatening complications and considerable costs [5]. Therefore, many patients with PVCs continue to live with persistent and debilitating symptoms.

Symptom preoccupation, which involves excessive attention to symptoms, fear of symptoms, and associated avoidance behavior, is related to increased symptom severity and low disease-specific quality of life (QoL) in other somatic conditions [6,7]. Studies of atrial fibrillation (AF) have shown that symptom preoccupation, rather than the objective arrhythmia burden, explains elevated symptom severity and impaired QoL [8,9]. However, few studies have investigated QoL in patients with PVCs [10], and the factors underpinning symptom severity and QoL in these patients are not well understood. Given the similar symptom presentation as that associated with AF, we hypothesized that symptom preoccupation is likely to play a role in the subjective symptom experience for patients with PVCs. We have defined symptom preoccupation in the context of cardiac arrhythmias (AF and PVCs) as the fear of experiencing and triggering cardiac-related symptoms, hypervigilance toward cardiac symptoms, persistent worry about complications, and arrhythmia-related avoidance of physical and social activities [11,12]

Cognitive behavioral therapy (CBT) has been found to be effective in reducing disability associated with both anxiety disorders [13] and somatic conditions where symptom preoccupation is prevalent [6,7,13]. In a series of clinical studies, Särnholm et al [11,12,14] developed AF-specific CBT (AF-CBT) targeting symptom preoccupation, which resulted in significant improvements in AF-specific QoL and self-reported symptom severity. Therefore, the purpose of this study was to further adapt and evaluate the AF-CBT protocol to target symptom preoccupation in patients with symptomatic idiopathic PVCs.

PVC-specific CBT (PVC-CBT) aims to break the cycle of cardiac-related fear and hypervigilance as well as arrhythmia-specific avoidance behavior and disability through repeated exposure to cardiac-related symptoms and situations that have been avoided due to fear, negative emotional responses, and apprehension of experiencing PVCs. We also aimed to investigate changes in the objective arrythmia burden as measured with electrocardiogram (ECG), along with the role of symptom preoccupation as a potential mediator of the treatment effect on PVC-specific QoL.

Methods

Study Design

In this uncontrolled pilot trial, we used a pretest-posttest design with 3- and 6-month follow-ups. Self-rated measures were completed online using a secure web-based assessment tool and were collected at pretreatment, post treatment, and at 3- and 6-month follow-ups. The primary outcome and potential mediators were also collected weekly during treatment. The aim was to include 30 participants, yielding a power of 80% to detect at least a moderate increase in the main outcome measure, corresponding to an effect size of Cohen d=0.65. However, due to time constraints, we made a pragmatic decision to stop recruitment at 19 participants. This decision was based on the clinical impression of substantial improvements in the first 15 treated participants and experiences from the previous pilot study of AF-CBT, where large average improvement was observed in 19 participants [11].

Participants

Participants were recruited nationwide from Sweden by advertisement in social media and newspapers. To be eligible for the study, participants had to fulfill the following inclusion criteria: (1) 18-70 years old, (2) diagnosed with PVCs with impairing or bothering PVC-associated cardiac symptoms, (3) on medication in accordance with current guidelines [5], and (4) able to read and write in Swedish. Participants were excluded if they had (1) any structural heart disease, including previous myocardial infarction, heart failure with preserved or reduced left ventricular ejection fraction, valvular disease, or previous cardiac surgery; (2) other arrhythmia or severe medical illness; (3) were scheduled for ablation therapy or any other cardiovascular intervention; (4) any medical restriction to physical exercise; (5) severe psychiatric disorder, severe depression, or risk of suicide; or (6) alcohol dependency. All participants underwent cardiac and psychological assessments to ensure that eligibility criteria were met. Participants were asked not to engage in other psychological treatment and only make necessary changes in medication during study participation. Participants were recruited and treated between January 2021 and October 2021, and the last 6-month follow-up was conducted in April 2022.

Ethical Considerations

The Regional Ethics Review Board in Gothenburg approved the trial protocol (Dnr 2020–05809) and the study was registered on ClinicalTrials.gov (NCT05087238). The study was conducted in accordance with the Declaration of Helsinki and all respondents provided informed consent prior to their involvement in the study, wherein they were given detailed information regarding the study's purpose, procedures, potential risks, benefits, and their rights as study participants. The informed consent form outlined these aspects clearly, and participants were given the opportunity to ask questions before agreeing to participate. All identifying data were stored on secure servers and data analysis was conducted on pseudonymized data. No compensation was given for participating. This study report adheres to the TREND (Transparent Reporting of Evaluations with Nonrandomized Designs) statement checklist for nonrandomized interventions [15]. The authors assure the completeness and accuracy of the data and adherence to the trial protocol.

Procedure

Applicants registered at the study's secure webpage and completed an online screening, including informed digital consent, demographic questions, and medical history. Applicants also completed the Alcohol Use Disorders Identification Test [16], the 9-item Patient Health Questionnaire (PHQ-9) measuring depressive symptoms [17], the Atrial Fibrillation Effects on Quality of Life (AFEQT) questionnaire adapted for PVCs (primary outcome measure), and the Cardiac Anxiety Questionnaire (CAQ) [18]. The cardiac study nurse (EÓ) then screened the applicants' medical records and conducted clinical telephone interviews to ensure that eligibility criteria were met. Eligible patients then underwent a structured telephone-based psychological assessment by a clinical psychologist. All clinical assessments and the cardiac parameters from the medical chart were reviewed by the study cardiologist (HS) before a decision on inclusion was made.

Intervention

The PVC-CBT intervention consisted of 10 weekly face-to-face sessions with a clinical psychologist (BEL) delivered via videoconference in conjunction with online text-based modules accessed through a secure web-based platform. All theoretical elements were presented verbally in sessions 1-4 and summarized in text-based form online together with homework assignments. This design allowed us to combine the flexibility of a face-to face session with the scaffolding structure of the internet-based format. The clinical psychologist could consult the study cardiologist in cases of uncertainties regarding participants' health. Due to technical shortcomings of the videoconference application, some of the weekly sessions were delivered by telephone and one participant received all sessions by telephone. The last 6 sessions focused on continuing working with the central elements of treatment presented in sessions 1-4. Participants were recommended to spend 30 minutes a day on the homework assignments. The treatment was based on the AF-specific CBT protocol [11,12,14] and adapted to patients with PVCs by authors BEL and JS.

The PVC-CBT was designed to target symptom preoccupation (ie, fear and hypervigilance toward cardiac-related symptoms and avoidance behavior) and included the following interventions: (1) education on PVCs and common psychological reactions to PVC symptoms, and the role of control and avoidance behavior in maintaining fear and hypervigilance of PVC symptoms; (2) interoceptive exposure to physical sensations similar to PVC symptoms by performing physical exercises such as increasing the heart rate and inducing palpitations by running on the spot or inducing dyspnea by excessive breathing to reduce the fear of symptoms and hypervigilance; (3) self-observation of cardiac symptoms, thoughts, feelings, and behavioral impulses to reduce fear and hypervigilance, serving as a form of interoceptive exposure technique; (4) in vivo exposure to avoided activities that were anticipated to elicit or potentially exacerbate PVC symptoms (such as vigorous exercise) or situations in which PVC symptoms are unwanted (such as engaging in leisure activities or driving); and (5) strategies on how to refrain from behaviors that serve to control symptoms, such as pulse checking, and how to handle worry when conducting exposure exercises. Participants were encouraged to combine interventions 2-5 to maximize the effect of exposure (ie, inducing dyspnea by excessive breathing [interoceptive exposure] before taking a walk alone in the woods [in vivo exposure]) and then using self-observation while experiencing symptoms instead of checking their pulse. Participants were also encouraged to view symptomatic episodes as opportunities to practice the skills acquired in treatment. The last module focused on (6) relapse prevention, including participants making their own plan for continuous practice of the acquired skills after the end of treatment. See Textbox 1 for an overview of the treatment.

Textbox 1. Overview of the treatment plan in cognitive behavioral therapy for patients with premature ventricular contractions (PVCs).

Education

- Education on PVCs
- The role of cardiac-related fear, hypervigilance, and behavior on symptoms and quality of life
- Self-observation of cardiac-related symptoms, thoughts, feelings, and behavioral impulses

Interoceptive exposure

- Exposure to physical sensations associated with PVC symptoms

In vivo exposure

- Gradual exposure to avoided situations or activities that patients fear may elicit or aggravate PVC symptoms or where symptoms are unwanted
- Combining in vivo exposure with interoceptive exposure while refraining from control and safety behavior

Relapse prevention

- Prevention of relapse into control or avoidance behavior by identifying risk situations

Assessments

Design

All outcome measures were completed online, with no interference of study personnel, at pretreatment, post treatment, and at 3- and 6-month follow-ups, except when noted otherwise.

Primary Outcome

In the absence of a validated PVC-specific QoL measure, we used the AFEQT [19] as the primary outcome measure and adapted the questionnaire to PVCs (AFEQT-PVC). The AFEQT is a well-validated instrument for assessing self-reported AF-specific QoL in four domains: AF symptoms, impact on physical and social activities, medical treatment concerns, and satisfaction with AF treatment. The scale consists of 20 items with a total score ranging from 0 (severe AF symptoms and disability) to 100 (no AF symptoms and disability) [19]. The AFEQT has been shown to be sensitive to clinical change, with a change of 18.9 points corresponding to a meaningful improvement as assessed by a physician [19,20]. The original structure of the AFEQT questionnaire was preserved, only removing the last 6 items measuring satisfaction and concern with current medical treatment as these issues were not targeted in the study treatment. All other items were retained except for changing the wording relevant to AF to be relevant to PVCs (eg, “On a scale of 1 to 7, over the past 4 weeks, as a result of your extra heart beats, *how* much did the feelings below bother you?”). The AFEQT-PVC measures PVC-related QoL in the following domains: arrhythmia symptoms and impairment in physical and social activities. The adapted scale consists of 16 items (0-5), with a total score ranging from 0 (severe symptoms and disability) to 100 (no symptoms or disability).

Secondary Outcomes

Secondary outcome measures included the Symptoms Checklist (SCL), which consists of two subscales measuring the frequency and severity of arrhythmia-specific symptoms [21], and the CAQ, which was used to measure symptom preoccupation and consists of three subscales: (1) cardiac-related fear, (2) attention to cardiac-related symptoms, and (3) cardiac-related avoidance [18]. QoL was measured with the 12-item Short-Form Health Survey (SF-12), which contains two subscales measuring physical health–related QoL (PCS-12) and mental health–related QoL (MCS-12) [22]. To assess fear toward bodily symptoms, the Body Sensations Questionnaire (BSQ) [23] was used. Stress reactivity was measured with the Perceived Stress Scale (PSS-4) [24] and physical activity was measured with the Godin-Shepard Leisure Time Physical Activity Questionnaire (GSLTPAQ) [25]. Depressive symptoms were measured with the PHQ-9 [17] and general anxiety was measured with the General Anxiety Disorder scale (GAD-7) [26]. Treatment credibility was measured with the Credibility/Expectancy Scale [27] in the second week of treatment.

Patient satisfaction with the treatment was assessed with the Client Satisfaction Questionnaire (CSQ) [28] post treatment. Potential adverse events from treatment were assessed at post treatment and at 3- and 6-month follow-ups. Participants were instructed to report and rate the short- and long-term discomfort caused by the adverse event from 0 (did not affect me at all) to 3 (affected me very negatively) [6].

ECG Measurements of Objective PVC Burden

To assess changes in the objective PVC burden (number of PVCs per day), participants wore a three-channel ambulatory ECG patch (ePatch [29]) continuously for 5 days at pretreatment, post treatment, and at the 6-month follow-up. All participants contributed with three recordings except one participant who missed the 6-month follow-up recording.

As a measure of the subjective experience of the PVC burden, participants were instructed to indicate symptoms by tapping a button on the patch recorder when experiencing symptoms of PVCs. The patch recorder was delivered via mail together with written instructions. The ECG data were analyzed by an experienced consultant using specialized software (Cardiologs; Cardiologs Technologies). The consultant was blinded to participant ID and assessment occasion.

Statistical Analysis

Linear mixed modeling was performed in Stata/IC 16.0 to analyze the change in the estimated mean for assessments from pretreatment to post treatment and from pretreatment to 3-month and 6-month follow-up, respectively. Effect sizes of within-group changes (Cohen *d*) were calculated as the mean change between the two compared time points (pretreatment to post treatment, pretreatment to 3-month follow-up, and pretreatment to 6-month follow-up) divided by each measure's standard deviation at baseline. The 95% CIs for effect sizes were calculated in R [30] using bootstrapping with 5000 samples. Data were analyzed in an intention-to-treat design, meaning that all participants were included in the analyses regardless of treatment completion status.

The change in objective PVC burden and in the self-reported PVC burden (ie, indicating PVC symptoms) was analyzed using the ECG measurements from pretreatment to post treatment and from pretreatment to 6-month follow-up in Stata/IC 16.0. Profile analysis with a Poisson generalized estimation equation model and log-link function was used for the incidence rate of objective PVCs and self-reported PVCs by tapping the patch recorder device.

Mediation Analysis

The potential mediating effects of symptom preoccupation (CAQ) on the effect of the treatment on the primary outcome measure were analyzed in an exploratory mediation analysis using the weekly version of the AFEQT-PVC, only differing from the AFEQT-PVC in that participants were asked to recall a period of 1 week instead of 1 month. The mediation analyses were conducted based on the 10 weekly measurements collected during the treatment (ie, at the beginning of the first treatment week until the beginning of the tenth treatment week).

The three subscales (attention, avoidance, and fear) as well as the total score of the CAQ were included as indicators of symptom preoccupation. To control for nonspecific improvement, we used the weekly versions of PSS-4 (perceived stress) and GSLTPAQ (physical activity), which are not targeted in the CBT treatment, as competing mediators. The analyses were performed in line with the procedure described by Baron and Kenny [31] and further developed by Preacher and Hayes [32]. With the purpose of investigating to what extent the changes in outcome could be explained by changes in the mediators, we conducted both single mediator analyses, in which each mediator was tested separately, and multiple mediator analyses, in which we included all mediators to compete. This design allowed us to study the relative contribution of each mediator to the improvement on the outcome weekly AFEQT-PVC.

We hypothesized a gradual and linear improvement in outcome during treatment with an effect of treatment week on mediators and the outcome. Further, we expected an association between the mediators and outcome during treatment. Both the single and multiple mediation analyses were performed in three steps. First, the association between treatment week and the mediator(s) (ie, a-path) was estimated. Second, while controlling for treatment week, the association between the mediator(s) and weekly AFEQT-PVC (ie, b-path) over the course of the therapy was estimated. Third, the ab-product, which is the indirect or mediated effect (ie, the contribution of the change in the mediator on the effect of treatment week on the outcome), was calculated by multiplying the a- and b-path estimates for each mediator. In the second set of analyses, which included all mediators, the first step was conducted separately for each mediator as the dependent variable and the second step included all mediators as independent variables. To account for dependency between the weekly measurements, all analyses were based on linear mixed models, with random intercept 95% CIs for the indirect effects (ie, the ab-products) estimated using 5000 bootstrap replications of all analyses; the criterion for a statistically significant mediation effect was that the 95% CI did not contain zero [32].

Results

Sample

Table 1 displays the characteristics of the participants, and Figure 1 illustrates the participant flow through the trial. The included sample (N=19) predominantly comprised women (14/19, 74%). The mean age was 50.0 (SD 15.2) years and the self-reported mean time since the PVC diagnosis was 5 (SD 5.3) years.

Table 1. Characteristics of study participants at baseline (N=19).

Characteristics	Value
Women, n (%)	14 (74)
Age (years), mean (SD)	50 (15)
Employment status, n (%)	
Employed	12 (63)
Retired	5 (26)
Self-employed	1 (5)
Student	1 (5)
Highest completed education, n (%)	
Secondary	4 (21)
Tertiary	15 (79)
PVC[a] duration (years), mean (SD)	5 (5)
Current medication, n (%)	
Beta-blockers	11 (58)
Calcium-channel blocker	2 (11)
Antiarrhythmics	1 (5)
ACEi[b]/ARB[c]	1 (5)
Anticoagulation	1 (5)
SSRI[d]	1 (5)
Thyroid replacement therapy	3 (16)
Medical disorders, n (%)	
Hypertension	2 (11)
Dyslipidemia	2 (11)
Obstructive sleep apnea	2 (11)
Hypothyroidism	4 (21)
Comorbid psychiatric conditions, n (%)	
Any psychiatric condition	13 (68)
Depressed mood	2 (11)
Excessive worry	4 (21)
Social anxiety	3 (16)
Panic attacks	2 (11)
Agoraphobia	2 (11)
Trauma-related stress symptoms	2 (11)
Exhaustion symptoms	1 (5)
Sleeping impairment	5 (26)
Previous psychological treatment, n (%)	13 (68)

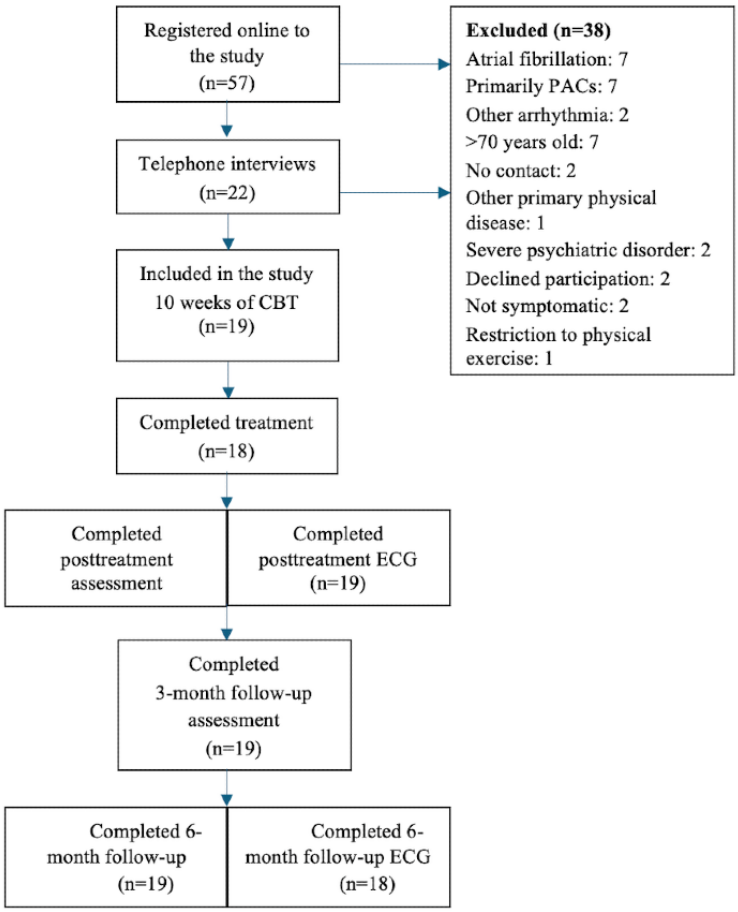
[a]PVC: premature ventricular contraction.

[b]ACEi: angiotensin converter-enzyme inhibitor.

[c]ARB: angiotensin receptor blocker.

[d]SSRI: selective serotonin reuptake inhibitor.

Figure 1. Flow of participants trough the trial. CBT: cognitive behavioral therapy; ECG: electrocardiogram; PAC: premature atrial contractions; PVC: premature ventricular contractions.



Treatment Activity

Mean session attendance was 9.6 (SD 1.9) sessions, ranging from 2 to 10 sessions. In total, 18 of the 19 participants (95%) were considered treatment completers, meaning that they completed at least 3 sessions and engaged in interoceptive and in vivo exposure exercises, and thus received the core components of the treatment. The one noncompleter attended 2 sessions. There were no missing self-assessment data at any assessment point, whereas one ECG measurement was missing at the 6-month follow-up.

Primary and Secondary Outcomes

Table 2 shows the scores for the continuous outcomes at all assessment points. We observed substantial improvements in PVC-specific QoL (AFEQT-PVC), with large within-group effect sizes at the pretreatment-to-posttreatment assessment. Furthermore, we observed large pretreatment-to-posttreatment reductions in the self-reported frequency (SCL frequency) and severity (SCL severity) of arrhythmia symptoms, and in symptom preoccupation as measured by the total CAQ score as well as on all three subscales of the CAQ: fear, avoidance, and attention. Large effect sizes were also observed for mental health–related QoL (MCS-12), bodily symptoms (BSQ), and depressive symptoms (PHQ-9). Moderate effect sizes were observed on self-perceived stress (PSS-4), general anxiety (GAD-7), and physical activity (GSLTPAQ). One participant (1/19, 5%) reported substantially higher levels of physical activity than the others, leading to a 7-fold increase in variance on the GSLTPAQ. This participant was deemed an outlier and removed from the outcome analysis. No significant effect was seen on the physical health–related QoL (PCS-12). Results for all measures were sustained at the 3- and 6-month follow-ups compared with baseline assessments, except for the PSS-4 score that was nonsignificant at 3-month follow-up.

Table 2. Continuous treatment outcome measures and mixed-effects regression model results.

Measure and assessment time point	Mean (SD)	Change from pretreatment	
		Cohen *d*[a] (95% CI)[b]	*P* value
AFEQT-PVC[c]			
Pretest	58.3 (13.8)	—[d]	—
Posttest	80.7 (16.1)	1.62 (1.07 to 2.28)	<.001
3-month follow-up	81 (17.5)	1.64 (1.06 to 2.36)	<.001
6-month follow-up	79.5 (19.4)	1.53 (0.92 to 2.18)	<.001
SC[e] frequency			
Pretest	19.6 (7.8)	—	—
Posttest	13 (7.2)	0.84 (0.47 to 1.2)	<.001
3-month follow-up	12.3 (7.1)	0.93 (0.51 to 1.33)	<.001
6-month follow-up	12.2 (8.0)	0.94 (0.52 to 1.4)	<.001
SCL severity			
Pretest	18.2 (6.4)	—	—
Posttest	10 (6.1)	1.14 (0.75 to 1.61)	<.001
3-month follow-up	10 (6.1)	1.29 (0.84 to 1.76)	<.001
6-month follow-up	9.9 (5.8)	1.30 (0.89 to 1.81)	<.001
CAQ[f]			
Pretest	35.6 (9.2)	—	—
Posttest	19.6 (9.9)	1.73 (1.27 to 2.38)	<.001
3-month follow-up	18.8 (0.5)	1.82 (1.26 to 2.39)	<.001
6-month follow-up	19.7 (12.9)	1.71 (1.03 to 2.54)	<.001
CAQ fear			
Pretest	16.5 (5.3)	—	—
Posttest	9.9 (5.3)	1.22 (0.78 to 1.71)	<.001
3-month follow-up	9.3 (5.3)	1.35 (0.94 to 1.87)	<.001
6-month follow-up	10.3 (6.4)	1.16 (0.67 to 1.78)	<.001
CAQ avoid			
Pretest	8.2 (4.9)	—	—
Posttest	3.9 (4.2)	0.86 (0.49 to 1.39)	<.001
3-month follow-up	3.7 (3.9)	0.91 (0.5 to 1.41)	<.001
6-month follow-up	3.7 (4.6)	0.92 (0.49 to 1.44)	<.001
CAQ attention			
Pretest	10.8 (2.5)	—	—
Posttest	5.7 (2.4)	2.05 (1.33 to 2.97)	<.001
3-month follow-up	5.8 (2.8)	2.03 (1.26 to 3.01)	<.001
6-month follow-up	5.8 (3.3)	2.03 (1.15 to 3.09)	<.001
MCS-12[g]			
Pretest	37.8 (9.8)	—	—
Posttest	47.9 (9.5)	1.03 (0.6 to 1.54)	<.001
3-month follow-up	47.0 (11)	0.94 (0.25 to 1.55)	<.001
6-month follow-up	47.9 (10.8)	1.03 (0.39 to 1.62)	<.001

Measure and assessment time point	Mean (SD)	Change from pretreatment	
		Cohen *d*[a] (95% CI)[b]	*P* value
PCS-12[h]			
Pretest	50.9 (8.3)	—	—
Posttest	52.5 (6.4)	0.18 (–0.4 to 0.5)	.26
3-month follow-up	50 (6.7)	0.11 (0.18 to 0.87)	.49
6-month follow-up	51 (7.5)	0.01 (–0.51 to 0.34)	.95
BSQ[i]			
Pretest	35.3 (8.9)	—	—
Posttest	28.0 (8.7)	0.81 (0.34 to 1.46)	<.001
3-month follow-up	26.0 (7)	1.04 (0.53 to 1.59)	<.001
6-month follow-up	27.1 (8.5)	0.91 (0.44 to 1.63)	<.001
PSS-4[j]			
Pretest	5.7 (3.2)	—	—
Posttest	4.2 (2.9)	0.50 (0.26 to 0.89)	.01
3-month follow-up	5.1 (3.5)	0.20 (–0.23 to 0.73)	.28
6-month follow-up	4.4 [m] (2.9)	0.43 (0 to 0.95)	.02
GLSTPAQ[k]			
Pretest	32.1 (24.1)	—	—
Posttest	49.3 (24.5)	0.71 (0.34 to 1.18)	.01
3-month follow-up	55.6 (42.2)	0.97 (0.49 to 2.28)	.001
6-month follow-up	53.4 (33.3)	0.88 (0.45 to 1.83)	.002
PHQ-9[l]			
Pretest	5.8 (2.8)	—	—
Posttest	3 (2.6)	1.03 (0.58 to 1.52)	<.001
3-month follow-up	3.8 (3.9)	0.71 (–0.27 to 1.32)	.02
6-month follow-up	3.8 (3.9)	0.73 (–0.06 to 1.34)	.01
GAD-7[m]			
Pretest	6.5 (3.2)	—	—
Posttest	4.2 (3.9)	0.73 (0.17 to 1.42)	.008
3-month follow-up	3.8 (3.4)	0.84 (0.24 to 1.47)	.003

Measure and assessment time point	Mean (SD)	Change from pretreatment	
		Cohen *d*[a] (95% CI)[b]	*P* value
6-month follow-up	3.3 (3.7)	1.02 (0.11 to 1.59)	<.001

[a]Within-group effect size.

[b]95% CIs are based on 5000 bootstrap replications.

[c]AFEQT-PVC: Atrial Fibrillation Effects on Quality of Life adapted for PVCs.

[d]Not applicable; all assessments at all other time points were compared to the pretest level.

[e]SCL: Symptoms Checklist.

[f]CAQ: Cardiac Anxiety Questionnaire.

[g]MCS-12: 12-Item Short-Form Survey mental health subscale.

[h]PCS-12: 12-Item Short-Form Survey physical health subscale.

[i]BSQ: Bodily Symptoms Questionnaire.

[j]PSS-4: Perceived Stress Scale.

[k]GLSTPAQ: Godin Shepard Leisure Time Physical Activity Questionnaire; one outlier was removed from this analysis.

[l]PHQ-9: Patient Health Questionnaire 9-item scale.

[m]GAD-7: Generalized Anxiety Disorder 7-item scale.

ECG Analyses

The ECG analyses (Table S1 in Multimedia Appendix 1) did not show any significant change in the objective PVC burden at post treatment (*P*=.87) or at 6-month follow-up (*P*=.21). However, we observed a statistically significant decrease in self-reported PVC symptoms (*P*=.006) as reported by tapping the patch-recorder symptom indicator when experiencing symptoms of PVCs at posttreatment assessment. These effects were sustained at 6-month follow-up (*P*=.003).

Mediation Analysis

The potential mediators and outcomes were measured weekly for 9 consecutive weeks during treatment. We collected a mean number of 17.7 observations per week out of 19 possible observations, and no less than 16 for any week (week 5). Table S2 in Multimedia Appendix 1 shows the estimated indirect effects (ab-products) and their 95% CIs for the three mediators when tested separately and in competition in a multiple mediator model.

In the single mediator analysis, the following measures had statistically significant ab-products: CAQ fear (0.64), CAQ attention (0.87), CAQ avoidance (1.01), and PSS-4 (0.21). This indicated a mediating effect of both symptom preoccupation and stress sensitivity on the main outcome AFEQT-PVC. The effect of physical activity based on the GSLTPAQ (–0.01) was not statistically significant, regardless of whether or not the outlying participant was included in the analysis. When allowing the measures to compete in explaining the change in outcome in AFEQT-PVC in a multiple mediator analysis, the ab-product of PSS-4 (0.12) was statistically significant but substantially lower than that of the attention (0.42) and avoidance (0.80) subscales of the CAQ, which also remained statistically significant. The fear subscale of the CAQ (0.14) was nonsignificant in the multiple mediator analysis.

Treatment Satisfaction

In total, 18 out of 19 participants (95%) reported that they were very satisfied (13/19, 68%) or satisfied (5/19, 26%) with the treatment. The mean score on the CSQ measure was 28.9 (SD 3.2) of a maximum 32 points, which indicated a high level of satisfaction with the treatment.

Changes in Medication and Cardiac Health

At post treatment, 4 of the 19 (21%) participants reported changes in their cardiac medication. Three participants quit using beta-blockers and one participant reported an increase in the use of beta-blockers. At the 6-month follow-up, 6 participants (32%) reported changes in cardiac medication; 1 (5%) quit using beta-blockers, 1 (5%) decreased the use of beta-blockers, and 4 (21%) increased their use of beta-blockers (including one participant who started using beta-blockers to treat hypertension rather than symptoms of PVCs). At 6-month follow-up, 1 (5%) participant had undergone invasive therapy (catheter ablation) and 1 (5%) participant reported slightly increased blood pressure. In addition, 1 (5%) participant had episodes of paroxysmal AF based on the analyses of the ECG data after the 6-month follow-up.

Adverse Events

At post treatment, among the 19 participants, 2 (11%) reported adverse events from engaging in the study. One of the participants (5%) reported increased stress and worry from participating in the study and rated the negative impact of this event as of mild severity (1 out of 3) at the time of the event as well as on the residual discomfort from the event. The other participant reported two adverse events during treatment consisting of elevated cardiac symptoms, rated as having the highest severity (3 out of 3) at the time of the events as well as on residual discomfort. No other adverse events were reported at the 3-month follow-up. At the 6-month follow-up, 1 (5%) participant reported an adverse event consisting of an episode of increased frequency of palpitations, rated as of medium severity at the time of the event (2 out of 3) and of mild severity on residual discomfort from the event (1 out of 3).

Discussion

Principal Findings

To our knowledge, this is the first study to evaluate the potential efficacy and feasibility of PVC-CBT for symptom preoccupation in patients with symptomatic idiopathic PVCs. Substantial improvements were found on the primary outcome measure AFEQT-PVC with respect to PVC-specific QoL and self-reported PVC symptoms. We also observed large reductions in cardiac-related fear and hypervigilance and avoidance behavior as measured by the CAQ. Furthermore, medium to large improvements were observed on all other secondary measures, except for the physical health domain of the SF-12, which remained unchanged. ECG analyses showed a significant reduction in self-reported PVC symptoms, although the objective PVC burden was unchanged. All posttreatment results were sustained at 3- and 6-month follow-ups, with high adherence and participant satisfaction and few adverse events reported. These results are comparable to the results for exposure-based AF-CBT delivered face to face [11] and via the internet [12,14] as well as for CBT for functional somatic disorders [6,7,33-35]. The exploratory mediation analyses indicated that symptom preoccupation had a mediating effect on the impact of PVC-CBT on PVC-specific QoL and self-reported PVC symptoms. These results are consistent with previous results of AF-CBT [12,14] and of CBT for other somatic conditions [36-38]. Sustained improvements in PVC-specific QoL (AFEQT-PVC) and self-reported PVC symptoms (SCL and tapping the ECG) were observed despite the lack of change in the objective PVC burden, providing support for further investigation on the proposed role of symptom preoccupation as a potential maintaining mechanism of symptom focus and impairment in PVCs.

The promising findings of this pilot study, with large improvements in the outcome measures, high treatment adherence, and reported satisfaction with treatment as well as the limited report of adverse events, encourage further research into the efficacy of CBT in patients with symptomatic PVCs. The use of videoconferencing and text-based material delivered online in this study enabled access to treatment for patients who do not otherwise benefit from current treatment regimens and for patients in rural areas. If the results of this study can be replicated in further studies, PVC-specific CBT may be made available to larger groups of patients, irrespective of their geographical location. In the future, as digital devices are increasingly used in the detection and follow-up of patients with arrhythmias [39], online CBT may be combined with other telemedicine applications as part of a remote management strategy in patients with PVC or other arrhythmic conditions.

Limitations

There are several limitations to this study, which should be considered when interpreting the results. The within-group design with a lack of control group precludes deriving a firm conclusion as to whether the results are true effects of the intervention or caused by extraneous variables such as the passage of time, expectancy of improvement, or attention from a caregiver. Another limitation is the reliance on self-reported measures, which could raise validity concerns due to the subjective nature of self-reporting. Nevertheless, the use of self-reported outcomes of QoL in arrythmia intervention research is endorsed [40]. All self-reported measures used in this study are well-validated, except for the main outcome measure (AFEQT-PVC), which has not been validated for the PVC population. Unfortunately, to our knowledge, there is no validated PVC-specific QoL questionnaire, and while the AFEQT is a well-validated measure for AF-specific QoL and patients with PVCs show a similar symptom presentation and arrhythmia-related disability to those of patients with AF, important aspects of QoL among patients with PVCs may not be adequately reflected by this measure. In addition, the ECG patch measurement period of 5 days may have been too short to measure the objective PVC burden accurately. A more reliable measure could have been obtained by using an implantable loop recorder. However, this was considered too intrusive for a secondary outcome measure. Inferences on generalizability are also limited by the small sample size and the skewed sex distribution with a majority of female participants (14/19, 74%). A possible selection bias may have been introduced due to recruitment mainly from advertisements on social media, which may affect the generalizability to patients in routine care. However, the treatment is designed for patients with symptom preoccupation who are willing to engage in psychological intervention and thus the selected recruitment approach may also have lowered the threshold for seeking treatment. A majority (13/19, 68%) of the participants had previous experience of psychological treatment, which may have made them more susceptible to the CBT treatment.

Conclusions

The main objective of this pilot study was to investigate the potential efficacy of PVC-CBT for patients with symptomatic idiopathic PVCs delivered by videoconference together with online text-based modules and homework assignments. The large improvements on PVC-specific QoL and the indication that the effect on PVC-specific QoL was mediated by reductions in symptom preoccupation suggest potential efficacy of exposure-based CBT targeting symptom preoccupation for this patient group. Randomized controlled trials are warranted to confirm our findings in larger patient samples.

Acknowledgments

Funding of this study was made possible by grants from Karolinska University Hospital. The funding organization did not exert any influence on the study design, execution, data analysis, or interpretation. All researchers were independent of the funders. The authors wish to express their gratitude to Professor Matteo Bottai at Karolinska Institutet for statistical consultation and analysis and to Elin Westberg for conducting the electrocardiogram analysis. Finally, we would like to thank the participants of the study for their time and effort.

Data Availability

The data sets generated during and/or analyzed during this study are not publicly available due to European Union law regulating sensitive personal data. Access to data requires approval from the Swedish Ethical Review Authority and data sharing agreement with Karolinska University Hospital.

Authors' Contributions

BEL, JS, BL, FB, and HS were responsible for study supervision. BEL, JS, and BL carried out the statistical analysis. BEL drafted the manuscript, which was revised by all authors.

Conflicts of Interest

FB declares personal fees for trial committee participation and lectures by Medtronic, Biotronic, Biosense Webster, Impulse Dynamics, Novartis, Orion, Boehringer, and Pfizer. BL has coauthored a self-help book based on exposure-based cognitive behavior therapy for health anxiety that is available in the public marketplace. The other authors have no conflicts to declare.

Multimedia Appendix 1
Detailed results of the electrocardiogram (ECG; Table S1) and exploratory mediation (Table S2) analyses.
[DOCX File , 23 KB - cardio_v8i1e53815_app1.docx]

References

1. von Rotz M, Aeschbacher S, Bossard M, Schoen T, Blum S, Schneider S, et al. Risk factors for premature ventricular contractions in young and healthy adults. Heart 2017 May 21;103(9):702-707. [doi: 10.1136/heartjnl-2016-309632] [Medline: 27798051]
2. Heckbert SR, Austin TR, Jensen PN, Floyd JS, Psaty BM, Soliman EZ, et al. Yield and consistency of arrhythmia detection with patch electrocardiographic monitoring: The Multi-Ethnic Study of Atherosclerosis. J Electrocardiol 2018 Nov;51(6):997-1002 [FREE Full text] [doi: 10.1016/j.jelectrocard.2018.07.027] [Medline: 30497763]
3. Marcus GM. Evaluation and management of premature ventricular complexes. Circulation 2020 Apr 28;141(17):1404-1418. [doi: 10.1161/CIRCULATIONAHA.119.042434] [Medline: 32339046]
4. Cronin E, Bogun FM, Maury P, Peichl P, Chen M, Namboodiri N, ESC Scientific Document Group. 2019 HRS/EHRA/APHRS/LAHRS Expert Consensus Statement on catheter ablation of ventricular arrhythmias. Europace 2019 Aug 01;21(8):1143-1144 [FREE Full text] [doi: 10.1093/europace/euz132] [Medline: 31075787]
5. Zeppenfeld K, Tfelt-Hansen J, de Riva M, Winkel BG, Behr ER, Blom NA, ESC Scientific Document Group. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. Eur Heart J 2022 Oct 21;43(40):3997-4126. [doi: 10.1093/eurheartj/ehac262] [Medline: 36017572]
6. Ljótsson B, Hesser H, Andersson E, Lackner JM, El Alaoui S, Falk L, et al. Provoking symptoms to relieve symptoms: a randomized controlled dismantling study of exposure therapy in irritable bowel syndrome. Behav Res Ther 2014 Apr;55:27-39. [doi: 10.1016/j.brat.2014.01.007] [Medline: 24584055]
7. Hedman-Lagerlöf M, Hedman-Lagerlöf E, Axelsson E, Ljótsson B, Engelbrektsson J, Hultkrantz S, et al. Internet-delivered exposure therapy for fibromyalgia: a randomized controlled trial. Clin J Pain 2018 Jun;34(6):532-542. [doi: 10.1097/AJP.0000000000000566] [Medline: 29077623]
8. Walters TE, Wick K, Tan G, Mearns M, Joseph SA, Morton JB, et al. Symptom severity and quality of life in patients with atrial fibrillation: psychological function outweighs clinical predictors. Int J Cardiol 2019 Mar 15;279:84-89. [doi: 10.1016/j.ijcard.2018.10.101] [Medline: 30442375]
9. Ong L, Cribbie R, Harris L, Dorian P, Newman D, Mangat I, et al. Psychological correlates of quality of life in atrial fibrillation. Qual Life Res 2006 Oct 7;15(8):1323-1333. [doi: 10.1007/s11136-006-0029-5] [Medline: 16826433]
10. Timmers L, Vervaet P, Goethals P, De Pooter J, Vanheuverswyn F, Van de Velde E, et al. Quality of life and symptoms in patients selected for ablation of ventricular extrasystoles: a prospective study. J Interv Card Electrophysiol 2022 Dec 02;65(3):601-607. [doi: 10.1007/s10840-021-01092-x] [Medline: 34855035]
11. Särnholm J, Skúladóttir H, Rück C, Pedersen SS, Braunschweig F, Ljótsson B. Exposure-based therapy for symptom preoccupation in atrial fibrillation: an uncontrolled pilot study. Behav Ther 2017 Nov;48(6):808-819. [doi: 10.1016/j.beth.2017.06.001] [Medline: 29029677]
12. Särnholm J, Skúladóttir H, Rück C, Klavebäck S, Ólafsdóttir E, Pedersen SS, et al. Internet-delivered exposure-based therapy for symptom preoccupation in atrial fibrillation: uncontrolled pilot trial. JMIR Cardio 2021 Mar 02;5(1):e24524 [FREE Full text] [doi: 10.2196/24524] [Medline: 33650972]
13. Craske MG, Treanor M, Conway CC, Zbozinek T, Vervliet B. Maximizing exposure therapy: an inhibitory learning approach. Behav Res Ther 2014 Jul;58:10-23 [FREE Full text] [doi: 10.1016/j.brat.2014.04.006] [Medline: 24864005]

14. Särnholm J, Skúladóttir H, Rück C, Axelsson E, Bonnert M, Bragesjö M, et al. Cognitive behavioral therapy improves quality of life in patients with symptomatic paroxysmal atrial fibrillation. J Am Coll Cardiol 2023 Jul 04;82(1):46-56. [doi: 10.1016/j.jacc.2023.04.044] [Medline: 37380303]
15. Des Jarlais DC, Lyles C, Crepaz N, TREND Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health 2004 Mar;94(3):361-366. [doi: 10.2105/ajph.94.3.361] [Medline: 14998794]
16. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun 24;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
17. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
18. Eifert GH, Thompson RN, Zvolensky MJ, Edwards K, Frazer NL, Haddad JW, et al. The Cardiac Anxiety Questionnaire: development and preliminary validity. Behav Res Ther 2000 Oct;38(10):1039-1053. [doi: 10.1016/s0005-7967(99)00132-1] [Medline: 11004742]
19. Spertus J, Dorian P, Bubien R, Lewis S, Godejohn D, Reynolds MR, et al. Development and validation of the Atrial Fibrillation Effect on QualiTy-of-Life (AFEQT) Questionnaire in patients with atrial fibrillation. Circ Arrhythm Electrophysiol 2011 Feb;4(1):15-25. [doi: 10.1161/CIRCEP.110.958033] [Medline: 21160035]
20. Dorian P, Burk C, Mullin CM, Bubien R, Godejohn D, Reynolds MR, et al. Interpreting changes in quality of life in atrial fibrillation: how much change is meaningful? Am Heart J 2013 Aug;166(2):381-387. [doi: 10.1016/j.ahj.2013.04.015] [Medline: 23895823]
21. Bubien RS, Knotts-Dolson SM, Plumb VJ, Kay GN. Effect of radiofrequency catheter ablation on health-related quality of life and activities of daily living in patients with recurrent arrhythmias. Circulation 1996 Oct 01;94(7):1585-1591. [doi: 10.1161/01.cir.94.7.1585] [Medline: 8840848]
22. Gandhi S, Salmon JW, Zhao SZ, Lambert BL, Gore PR, Conrad K. Psychometric evaluation of the 12-item short-form health survey (SF-12) in osteoarthritis and rheumatoid arthritis clinical trials. Clin Ther 2001 Jul;23(7):1080-1098. [doi: 10.1016/s0149-2918(01)80093-x] [Medline: 11519772]
23. Chambless DL, Caputo GC, Bright P, Gallagher R. Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. J Consult Clin Psychol 1984 Dec;52(6):1090-1097. [doi: 10.1037//0022-006x.52.6.1090] [Medline: 6520279]
24. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385. [doi: 10.2307/2136404]
25. Amireault S, Godin G. The Godin-Shephard Leisure-Time Physical Activity Questionnaire: validity evidence supporting its use for classifying healthy adults into active and insufficiently active categories. Percept Mot Skills 2015 Apr 01;120(2):604-622. [doi: 10.2466/03.27.pms.120v19x7]
26. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
27. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000 Jun;31(2):73-86. [doi: 10.1016/s0005-7916(00)00012-4] [Medline: 11132119]
28. Attkisson C, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann 1982 Jan;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x] [Medline: 10259963]
29. Saadi DB, Sørensen HBD, Hansen IH, Egstrup K, Jennum K, Hoppe K. ePatch - A Clinical Overview. DTU Orbit. URL: https://backend.orbit.dtu.dk/ws/portalfiles/portal/102966188/ePatch_A_Clinical_Overview_DTU_Technical_Report.pdf [accessed 2024-04-25]
30. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.r-project.org/ [accessed 2024-04-25]
31. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986 Dec;51(6):1173-1182. [doi: 10.1037//0022-3514.51.6.1173] [Medline: 3806354]
32. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods 2008 Aug;40(3):879-891. [doi: 10.3758/brm.40.3.879] [Medline: 18697684]
33. Bonnert M, Olén O, Lalouni M, Benninga MA, Bottai M, Engelbrektsson J, et al. Internet-delivered cognitive behavior therapy for adolescents with irritable bowel syndrome: a randomized controlled trial. Am J Gastroenterol 2017 Jan;112(1):152-162 [FREE Full text] [doi: 10.1038/ajg.2016.503] [Medline: 27845338]
34. Lalouni M, Ljótsson B, Bonnert M, Ssegonja R, Benninga M, Bjureberg J, et al. Clinical and cost effectiveness of online cognitive behavioral therapy in children with functional abdominal pain disorders. Clin Gastroenterol Hepatol 2019 Oct;17(11):2236-2244 [FREE Full text] [doi: 10.1016/j.cgh.2018.11.043] [Medline: 30502501]

35. Bonnert M, Olén O, Lalouni M, Hedman-Lagerlöf E, Särnholm J, Serlachius E, et al. Internet-delivered exposure-based cognitive-behavioral therapy for adolescents with functional abdominal pain or functional dyspepsia: a feasibility study. Behav Ther 2019 Jan;50(1):177-188 [FREE Full text] [doi: 10.1016/j.beth.2018.05.002] [Medline: 30661558]
36. Ljótsson B, Hesser H, Andersson E, Lindfors P, Hursti T, Rück C, et al. Mechanisms of change in an exposure-based treatment for irritable bowel syndrome. J Consult Clin Psychol 2013 Dec;81(6):1113-1126. [doi: 10.1037/a0033439] [Medline: 23750460]
37. Bonnert M, Olén O, Bjureberg J, Lalouni M, Hedman-Lagerlöf E, Serlachius E, et al. The role of avoidance behavior in the treatment of adolescents with irritable bowel syndrome: A mediation analysis. Behav Res Ther 2018 Jun;105:27-35 [FREE Full text] [doi: 10.1016/j.brat.2018.03.006] [Medline: 29614378]
38. Hedman-Lagerlöf M, Andersson E, Hedman-Lagerlöf E, Wicksell RK, Flink I, Ljótsson B. Approach as a key for success: reduced avoidance behaviour mediates the effect of exposure therapy for fibromyalgia. Behav Res Ther 2019 Nov;122:103478. [doi: 10.1016/j.brat.2019.103478] [Medline: 31585343]
39. Svennberg E, Tjong F, Goette A, Akoum N, Di Biase L, Bordachar P, et al. How to use digital devices to detect and manage arrhythmias: an EHRA practical guide. Europace 2022 Jul 15;24(6):979-1005. [doi: 10.1093/europace/euac038] [Medline: 35368065]
40. Steinberg BA, Dorian P, Anstrom KJ, Hess R, Mark DB, Noseworthy PA, et al. Patient-reported outcomes in atrial fibrillation research: results of a Clinicaltrials.gov analysis. JACC Clin Electrophysiol 2019 May;5(5):599-605 [FREE Full text] [doi: 10.1016/j.jacep.2019.03.008] [Medline: 31122382]

Abbreviations

AF: atrial fibrillation
AF-CBT: atrial fibrillation–specific cognitive behavioral therapy
AFEQT: Atrial Fibrillation Effects on Quality of Life
AFEQT-PVC: Atrial Fibrillation Effects on Quality of Life adapted to premature ventricular contraction
BSQ: Body Sensations Questionnaire
CAQ: Cardiac Anxiety Questionnaire
CBT: cognitive behavioral therapy
CSQ: Client Satisfaction Questionnaire
ECG: electrocardiogram
GAD-7: Generalized Anxiety Disorder scale
GSLTPAQ: Godin-Shepard Leisure Time Physical Activity Questionnaire
MCS-12: mental health–related quality of life subscale of the Short-Form Health Survey
PCS-12: physical health–related quality of life subscale of the Short-Form Health Survey
PHQ-9: 9-item Patient Health Questionnaire
PSS-4: Perceived Stress Scale
PVC: premature ventricular contraction
PVC-CBT: premature ventricular contraction–specific cognitive behavioral therapy
QoL: quality of life
SCL: Symptoms Checklist
SF-12: Short-Form Health Survey
TREND: Transparent Reporting of Evaluations with Nonrandomized Designs

Edited by A Mavragani; submitted 20.10.23; peer-reviewed by A AL-Asadi, IV George, M Hayıroğlu; comments to author 16.03.24; revised version received 05.04.24; accepted 09.04.24; published 07.05.24.

Please cite as:
Liliequist BE, Särnholm J, Skúladóttir H, Ólafsdóttir E, Ljótsson B, Braunschweig F
Cognitive Behavioral Therapy for Symptom Preoccupation Among Patients With Premature Ventricular Contractions: Nonrandomized Pretest-Posttest Study
JMIR Cardio 2024;8:e53815
URL: https://cardio.jmir.org/2024/1/e53815
doi:10.2196/53815
PMID:38713500

Original Paper

Factors That Influence Patient Satisfaction With the Service Quality of Home-Based Teleconsultation During the COVID-19 Pandemic: Cross-Sectional Survey Study

Guangxia Meng[1], MSN; Carrie McAiney[1], PhD; Ian McKillop[1], PhD; Christopher M Perlman[1], PhD; Shu-Feng Tsao[1], PhD; Helen Chen[1], PhD

School of Public Health Sciences, University of Walterloo, Waterloo, ON, Canada

Corresponding Author:
Helen Chen, PhD
School of Public Health Sciences
University of Walterloo
200 University Ave W
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567 ext 42131
Email: helen.chen@uwaterloo.ca

Abstract

Background: Ontario stroke prevention clinics primarily held in-person visits before the COVID-19 pandemic and then had to shift to a home-based teleconsultation delivery model using telephone or video to provide services during the pandemic. This change may have affected service quality and patient experiences.

Objective: This study seeks to understand patient satisfaction with Ontario stroke prevention clinics' rapid shift to a home-based teleconsultation delivery model used during the COVID-19 pandemic. The research question explores explanatory factors affecting patient satisfaction.

Methods: Using a cross-sectional service performance model, we surveyed patients who received telephone or video consultations at 2 Ontario stroke prevention clinics in 2021. This survey included closed- and open-ended questions. We used logistic regression and qualitative content analysis to understand factors affecting patient satisfaction with the quality of home-based teleconsultation services.

Results: The overall response rate to the web survey was 37.2% (128/344). The quantitative analysis was based on 110 responses, whereas the qualitative analysis included 97 responses. Logistic regression results revealed that responsiveness (adjusted odds ratio [AOR] 0.034, 95% CI 0.006-0.188; *P*<.001) and empathy (AOR 0.116, 95% CI 0.017-0.800; *P*=.03) were significant factors negatively associated with low satisfaction (scores of 1, 2, or 3 out of 5). The only characteristic positively associated with low satisfaction was when survey consent was provided by the substitute decision maker (AOR 6.592, 95% CI 1.452-29.927; *P*=.02). In the qualitative content analysis, patients with both low and high global satisfaction scores shared the same factors of service dissatisfaction (assurance, reliability, and empathy). The main subcategories associated with dissatisfaction were missing clinical activities, inadequate communication, administrative process issues, and absence of personal connection. Conversely, the high-satisfaction group offered more positive feedback on assurance, reliability, and empathy, as well as on having a competent clinician, appropriate patient selection, and excellent communication and empathy skills.

Conclusions: The insights gained from this study can be considered when designing home-based teleconsultation services to enhance patient experiences in stroke prevention care.

(JMIR Cardio 2024;8:e51439) doi:10.2196/51439

KEYWORDS

teleconsultation; secondary stroke prevention; telemedicine; service quality; patient satisfaction

Introduction

Secondary Stroke Prevention and Ontario Stroke Prevention Clinics

As of 2019, stroke was the second leading cause of disability worldwide for people aged >50 years, and it was the fourth leading cause of death in Canada from 2017 to 2019 [1,2]. The 36% decline in stroke mortality from 1990 to 2016 can be attributed to better prevention and management of stroke risks [3]. The INTERSTROKE study found that 90% of strokes are preventable owing to modifiable risk factors, including disease-related and behavioral lifestyle factors [4]. Secondary stroke prevention is crucial, as there is up to a 10% risk of recurrent stroke within 90 days of a transient ischemic attack (TIA) or minor stroke [5].

Approximately 80% of patients with minor stroke discharged from the emergency department in Ontario are referred to stroke prevention services [6]. Stroke prevention clinics provide rapid assessments, diagnostic tests, treatments, prevention, and education to reduce the risk of recurrent stroke [7]. Ontario's 41 stroke prevention clinics are integral to publicly funded health systems [7]. Stroke prevention clinic services are associated with a 25% reduction in mortality [8].

Before the COVID-19 pandemic, stroke prevention care in Ontario was predominantly delivered through in-person consultations. A small percentage of rural and northern Ontario stroke prevention clinics used teleconsultation at local satellite clinics to address access challenges [9]. One stroke prevention clinic conducted a pilot project offering follow-up home video visits from August 2018 to September 2019 [10]. The video consultation produced higher patient satisfaction, was considered safe by physicians, and was shown to be cost-effective in reducing health care costs and patient expenses [10].

In this study, home-based teleconsultation was defined as a synchronous consultation between a clinical service provider and a patient in their home to provide diagnostic or therapeutic advice through telephone or videoconference [11]. Despite a handful of cases in which home-based teleconsultation was used for follow-up care, before the COVID-19 pandemic [10], Ontario stroke prevention clinics had never used home-based teleconsultation to conduct synchronous, interactive, in-home patient visits for new referrals. A survey of >3000 Canadians with stroke, heart disease, or vascular impairment conducted in the spring of 2021 showed that 80% of respondents had had a teleconsultation during the pandemic [12]. The effect of the rapid change from in-person visits to home-based teleconsultation during the COVID-19 pandemic on patients' experiences was unknown. Patients with stroke are often older adults with multiple chronic conditions [13]. Older adults have lower telehealth service use overall [14]. The impact of service mode change on the older population of stroke prevention clinics needs further exploration.

Service Quality and Patient Satisfaction

Although the rapid transition to home-based teleconsultation may be a temporary response to the COVID-19 pandemic, it offers a significant opportunity to examine the service quality of home-based teleconsultation in stroke prevention clinics. Delivering safe, high-quality health services is the primary goal of health systems [15]. The literature's definition of health care service quality is commonly described as including 2 aspects. One views health care service quality as characteristics and features that meet clinicians' predetermined specifications and standards (such as professional or ethical standards); the other views it as characteristics and features that meet or exceed patients' needs and expectations [16]. Patients often cannot accurately assess the internal service quality as they lack the medical knowledge to judge [17]; however, patient satisfaction with a medical service is the primary determinant of service quality [18]. Patient satisfaction is essential for and meaningful to delivering high-quality care [19].

Patient satisfaction is generally regarded as patients' perception of care delivery as well as how their health needs have been addressed [20,21]. Patient satisfaction can be examined using direct and indirect indicators [21]. First, service quality can be measured by directly asking patients to rate their satisfaction with service quality via, for example, a single item with response options ranging from *very dissatisfied* to *very satisfied* [21]. However, the shortfall of single-item measurement is that we can not evaluate or identify a specific aspect of service quality [21]. The alternative approach is to ask patients to rate their experience of different aspects of care, but this indirect measure has the weakness of preemptive assumptions about the determinants of service quality [21]. To obtain an accurate measurement of patient satisfaction, we applied direct and indirect measurements. We asked one question on global satisfaction and applied a theory-guided questionnaire to assess patient satisfaction.

Service Performance Model

As health care quality is multidimensional, we chose an appropriate service quality model (SERVQUAL) to assess patient satisfaction. Examples of existing health care SERVQUALs include the SERVQUAL [22] and its derivative service performance model (SERVPERF) [23], Total Quality Management [24], Health Quality Model [25], service quality for a public hospital [26], and hospital quality model [27]. The Health Quality Model, service quality for a public hospital, and hospital quality model are derivatives of the SERVQUAL model and assess the care quality of hospital inpatient services. The SERVQUAL and SERVPERF models include 5 dimensions: tangibles, reliability, responsiveness, assurance, and empathy [22]. SERVQUAL tends to measure the difference between one's expectations and the actual performance of the service [22]. SERVPERF only focuses on the actual performance, and studies have shown that service quality is appropriately modeled using SERVPERF as an antecedent of satisfaction [23,28]. There are 22 service attributes listed within the 5 service dimensions of the SERVPERF model [29]. Textbox 1 presents each dimension and item in detail.

Textbox 1. Dimensions and items of the service performance model (adapted from Zeithaml et al [29]).

Tangibles: facilities, equipment, and the presence of personnel

- Up-to-date equipment
- Visually appealing physical facilities
- Neat-appearing employees
- Visually appealing materials associated with the service

Reliability: ability to perform the promised service responsibly and accurately

- The company keeps its promises to do something by a certain time
- The company shows a sincere interest in solving the customer's problem
- The company performs the service right the first time
- The company provides its services at the time it promises to do so
- The company insists on error-free records

Responsiveness: willingness to provide help and a prompt service to customers

- Employees of the company tell customers exactly when services will be performed
- Employees of the company provide a prompt service to customers
- Employees of the company are always willing to help customers
- Employees of the company are never too busy to respond to customers

Assurance: the knowledge and courtesy of employees and their ability toinspire trust and confidence

- Thebehaviorof employees of thecompany instills confidence incustomers
- Customers of the company feel safe in their transactions
- Employees of the company are consistently courteous with customers
- Employees of the company have the knowledge to answer customers'questions

Empathy:caring and understanding, whicha company provides or offers its customers in terms of its individualized and personalized attention

- The company gives customers individual attention
- The company has operating hours convenient to all itscustomers
- Employees of the company give customers personal attention
- The company has the customers' best interests at heart
- The employees of the company understand the specific needs of their customers

Study Aims and Research Question

Any new service model implementation should usually be well planned to improve user satisfaction; however, home-based teleconsultation at stroke prevention clinics was implemented without the usual planning. The rapid implementation could affect patients' experiences. As a result, it is vital to evaluate patient satisfaction with the quality of the teleconsultation service they received during the COVID-19 pandemic by assessing satisfaction with various service dimensions to identify aspects of service quality that patients are and are not satisfied with. This study aimed to explore the factors affecting patient satisfaction. The research question was as follows: "what are the patient-identified factors influencing patients' satisfaction with service quality in stroke prevention clinics' home-based teleconsultation service?"

Methods

Participants and Procedure

We conducted a web-based or telephone survey of patients who had at least one home-based teleconsultation, either the initial or follow-up visit, at the stroke prevention clinics. The study sites were 2 stroke prevention clinics at 2 tertiary hospitals in Ontario, Canada. The study sample consisted of individuals who received at least one home-based teleconsultation at a stroke prevention clinic between January 1, 2021, and November 30, 2021. A convenience sampling technique was used as we invited patients who lived in their homes and self-participated in the stroke prevention clinic home-based teleconsultation service during the COVID-19 pandemic. Our exclusion criteria included (1) patients who lived in a long-term care home or group home and (2) patients with dementia who could not participate in home-based teleconsultations.

To minimize volunteer bias and increase the response rate, we applied various data collection techniques to capture participants with and without internet access. A web survey was used for patients with email addresses, whereas a telephone survey was used for patients without email addresses. The questionnaire was administered between May 17, 2021, and December 10, 2021.

We developed a telephone script for recruitment to explain the research project in lay terms. In total, 2 modified-duty nurses from one site and neurologists from another site who were not part of the research team obtained permission from patients to be contacted by the research team. The list of email addresses or telephone numbers of patients who gave permission was shared with the research team. Participants who chose a web-based survey received an email with a brief cover letter explaining the study's purpose and their rights as study participants. Informed consent was via the web before accessing the questionnaire, and they were asked to click a box indicating that they agreed to complete the survey (Multimedia Appendix 1). Participants who chose a telephone survey received a mail-in cover letter, a consent form, and a copy of the survey (Multimedia Appendix 2).

Ethical Considerations

This study was reviewed by the research ethics boards of Southlake Regional Health Center and Mackenzie Health and was considered a continuous quality improvement project; thus, a full research ethics review was not required. This study was also reviewed by the University of Waterloo Office of Research Ethics (ORE 42686) and received ethical clearance.

Measures

Our study used SERVPERF to design a Likert-scale survey to assess patient satisfaction with home-based teleconsultation service quality in stroke prevention clinics. We acknowledge that patient satisfaction is subjective, with many determinants that may not be related to the SERVPERF model. The literature has indicated that patient satisfaction can be influenced by patient knowledge and expectations; therefore, other factors such as demographics (eg, age, gender, and education), clinical factors (eg, comorbidities, diagnosis, and number of visits), and experiences with teleconsultation can influence patient expectations [20,21]. We also included these factors in our survey. By considering other factors and applying direct and indirect measurements, we attempted to explore patient satisfaction using a holistic approach. The 18-item questionnaire used in our study was developed by referencing the telehealth service quality questionnaire developed by Yin et al [30] (see Multimedia Appendix 3 for a description).

The survey consisted of three components: (1) demographic, clinical, and telemedicine questions; (2) an 18-item Likert scale–based questionnaire measured on a 5-point scale, with 1 for *strongly disagree* and 5 for *strongly agree*; and (3) 6 open-ended follow-up questions (Multimedia Appendix 4). The demographic, clinical, and telemedicine independent variables were selected based on previous evidence from a literature review and clinical significance from a practice point of view [31]. We conducted a pilot study in March 2021 with 10 participants who had home-based teleconsultation from October 2020 to December 2020 and asked 6 additional questions about the survey content (Multimedia Appendix 5). Overall, patients were satisfied with the language and content of the survey, indicated in their feedback.

Data Collection

The web survey was conducted through a secured, password-protected REDCap (Research Electronic Data Capture; Vanderbilt University) website hosted at the University of Waterloo that supports research data collection [32]. Skype for Business (Skype Technologies) from the University of Waterloo, with recording and transcription functions, was set up for the research assistant to conduct the telephone survey.

Statistical Analysis

We applied quantitative and qualitative analysis to understand patient satisfaction. A binary outcome variable was defined as (1) a low-satisfaction group if the participants chose *very unsatisfied*, *dissatisfied*, and *neither satisfied nor satisfied* with the overall home-based teleconsultation service quality; and (2) a high-satisfaction group if the participants chose *satisfied* and *very satisfied*. We used SPSS for Windows (version 28.0.1; IBM Corp) for statistical analysis [33]. The Likert-scale questions were converted to numerical values. Using the item means, we generated each SERVPERF dimension score and an overall questionnaire score for each respondent's survey. There were 10 demographic, 7 clinical, and 6 technical-related independent variables (see Multimedia Appendix 6 for the definitions). Chi-square tests were used to identify the statistical significance between the categorical independent and binary outcome variables. The point biserial correlation was calculated to identify the correlation between a continuous independent variable and the binary outcome variable. To test the internal reliability of our instrument, we calculated the Cronbach α. As we had a large number of independent variables under consideration, a forward selection model was most suitable [34]. We used statistically significant variables correlating to the binary outcome variable in the stepwise binary logistic regression model, with $P<.05$ considered significant.

We used NVivo (QSR International), a software developed to organize and support the analysis of qualitative data. GM coded the entire data set, and ST independently coded 10 random samples. The results were compared and reached an initial 87.1% agreement. Discrepancies were discussed, and conflicts were resolved after further clarification of the definition of the codes. We applied direct content analysis to understand the service quality of the teleconsultation under study [35]. We used the 5 service dimensions and their operational definitions as the initial coding categories [36]. Next, we read each transcript and identified and categorized all the text that appeared to represent the operational definition of the code [35]. Text not categorized using the initial coding scheme would be considered for a new code. We summarized the categories of the entire data set and then divided them into low- and high-satisfaction groups to explore positive and negative patient perceptions. We compared the differences between the 2 groups that could explain the quantitative analysis results [37].

Results

Participant Characteristics

The response rate was 35.9% (104/290) for the web-based survey and 44% (24/54) for the telephone survey. A total of 110 (n=86, 78.2% web and n=24, 21.8% telephone) surveys were included for quantitative analysis, and 97 (n=74, 76% web and n=23, 24% telephone) surveys were included for direct content analysis. Figures 1 and 2 show a flowchart summarizing the subsequent exclusion of cases from the original number participants to arrive at the final analysis. A total of 97.3% (107/110) of the participants used telephone consultations. The percentages of missing values (1% to 8%) for each Likert-scale question were insignificant (<20%); therefore, the mean of each item was used to replace the missing data [38].

Figure 1. Sample flowchart for the web-based survey.

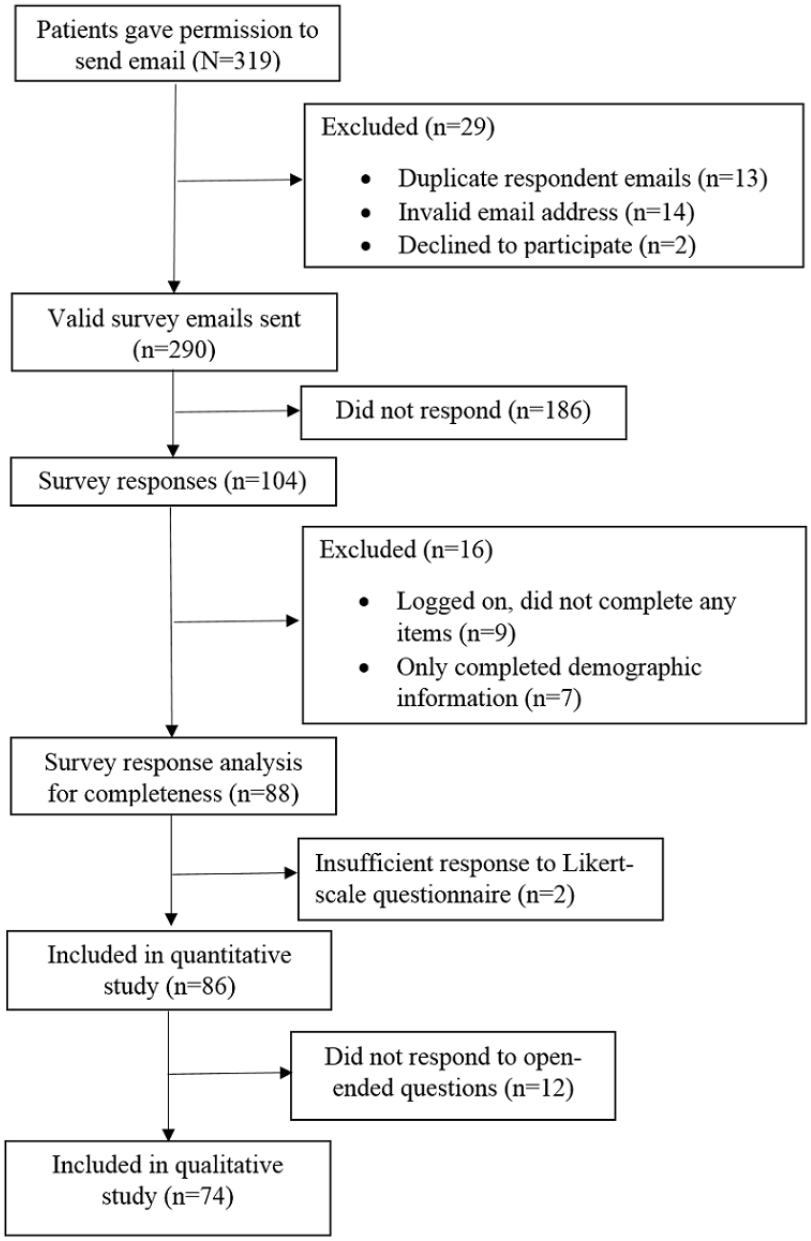
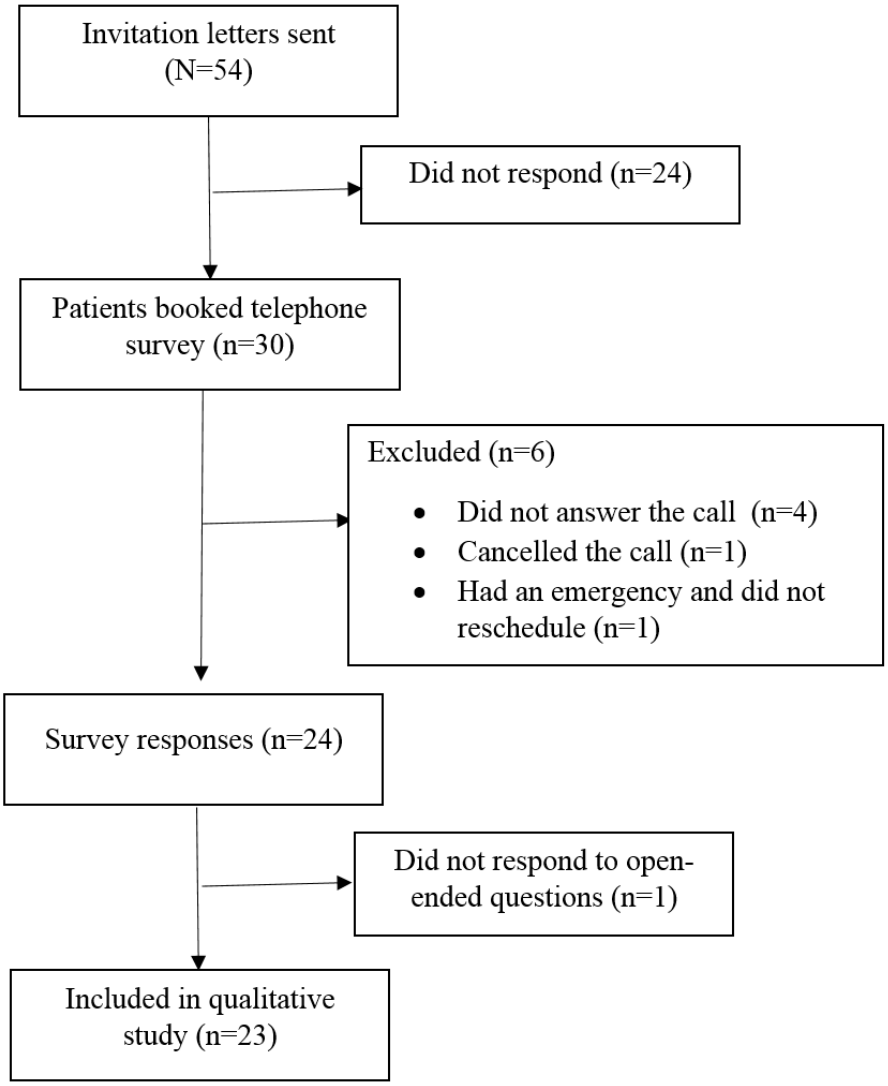


Figure 2. Sample flowchart for the telephone survey.



The descriptive statistics of the demographic, clinical, and telemedicine variables are presented in Tables 1-3. Briefly, most of the participants (99/110, 90%) were aged ≥55 years, were retired (80/109, 73.4%) and married (76/110, 69.1%), lived with others (85/109, 78%), and lived within 20 km of where the stroke prevention clinic was located (77/110, 70%). Only a few participants (7/109, 6.4%) had an educational level lower than high school, and most (67/109, 61.5%) had a postsecondary education (Table 1). Regarding clinical factors, most participants had a stroke diagnosis (71/110, 64.5%) and self-identified as having only one stroke risk factor (74/110, 67.3%). Most patients were new (101/110, 91.8%) to the stroke prevention clinics (Table 2). Many participants had relatively less experience with health technology. Although most of them owned digital equipment for teleconsultation (79/109, 72.5%), many of them had never used patient portals (94/108, 87%) and telemedicine (92/108, 85.2%) before the COVID-19 pandemic (Table 3).

Table 1. Demographic characteristics of the patients included in the study (N=110).

Variable	Total[a]	Low satisfaction (n=26)	High satisfaction (n=84)	*P* value
Age group (y), n (%)				.04[b]
<55	11 (10)	2 (7.7)	9 (10.7)	
55-64	18 (16.4)	3 (11.5)	15 (17.8)	
65-74	25 (25.5)	4 (15.4)	24 (28.6)	
75-84	40 (36.4)	16 (61.5)	24 (28.6)	
≥85	13 (11.8)	1 (3.8)	12 (14.3)	
Sex, n (%)				.52
Female	49 (44.5)	13 (50)	36 (42.9)	
Male	61 (55.5)	13 (50)	48 (57.1)	
Distance (km), mean (SD)	17.73 (16.52)	13.53 (11.65)	19.03 (17.61)	.31
1-20, n (%)	77 (70)	21 (80.8)	56 (66.7)	
21-40, n (%)	21 (19.1)	4 (15.4)	17 (20.2)	
>40, n (%)	12 (10.9)	1 (3.8)	10 (13.1)	
Education, n (%)				.52
Grade 8 or lower	7 (6.4)	0 (0)	7 (8.3)	
High school	35 (32.1)	10 (40)	25 (29.8)	
College	32 (29.4)	6 (24)	26 (31)	
University	24 (21.8)	6 (24)	18 (21.4)	
Graduate school	11 (10)	3 (12)	8 (9.5)	
Employment, n (%)				.68
Retired	80 (73.4)	20 (76.9)	60 (72.3)	
Working	17 (15.6)	2 (7.7)	15 (18.1)	
Unemployed	4 (3.7)	1 (3.8)	3 (3.6)	
Self-employed	3 (2.8)	1 (3.8)	2 (2.4)	
On disability	5 (4.6)	2 (7.7)	3 (3.6)	
Marital status, n (%)				.42
Married	76 (69.1)	18 (69.2)	58 (69.1)	
Divorced	11 (10)	4 (15.4)	7 (8.3)	
Widowed	17 (15.5)	4 (15.4)	13 (15.5)	
Single	6 (5.5)	0 (0)	6 (7.1)	
Living arrangement, n (%)				.69
With others	85 (78)	21 (80.8)	64 (77.1)	
Alone	24 (22)	5 (19.2)	19 (22.9)	
Transportation, n (%)				.24
Drives	67 (65)	12 (54.5)	55 (67.9)	
Relies on others	36 (35)	10 (45.5)	26 (32.1)	
Use of a cane or walker (yes), n (%)	30 (27.5)	8 (30.8)	22 (26.2)	.57
Language barrier (yes), n (%)	10 (9.1)	6 (23.1)	4 (4.8)	.005[c]
Hearing impaired (yes), n (%)	28 (25.5)	9 (34.6)	19 (22.6)	.22
Affecting phone conversations (yes)	14 (50)	6 (66.7)	8 (42.1)	
Vision impaired (yes), n (%)	33 (30)	8 (30.8)	25 (29.8)	.92

Variable	Total[a]	Low satisfaction (n=26)	High satisfaction (n=84)	*P* value
Affecting the use of a screen (yes)	16 (50)	5 (62.5)	11 (45.8)	.41
The survey was consented to by the SDM[d] (yes), n (%)	53 (49.1)	18 (69.2)	35 (41.7)	.01[b]
Web survey (yes), n (%)	86 (78.2)	24 (92.3)	62 (73.8)	.046[b]

[a]Note that the percentages are based on denominators that vary from the overall sample size of 110 because of missing data.
[b]*P*<.05.
[c]*P*<.01.
[d]SDM: substitute decision maker. Indicates that the survey was consented to and answered with the help of an SDM.

Table 2. Clinical characteristics of the patients included in the study (N=110).

Variable	Total, n (%)	Low satisfaction (n=26), n (%)	High satisfaction (n=84), n (%)	*P* value
Had stroke diagnosis	71 (64.5)	17 (65.4)	54 (64.3)	.92
Had stroke residual deficits	27 (38)[a]	9 (52.9)[b]	18 (33.3)[c]	.15
New referral	101 (91.8)	25 (96.2)	76 (90.5)	.36
Number of stroke risk factors				.40
0	14 (12.7)	3 (11.5)	11 (13.1)	
1	74 (67.3)	21 (80.8)	53 (63.1)	
2-4	19 (17.3)	1 (3.8)	18 (21.4)	
5-6	3 (2.7)	1 (3.8)	2 (2.4)	

[a]N=71.
[b]N=17.
[c]N=54.

Table 3. Telemedicine-related characteristics of the patients included in the study (N=110).

Variable	Total, n (%)[a]	Low satisfaction (n=26), n (%)	High satisfaction (n=84), n (%)	*P* value
Used telephone	107 (98.2)	26 (100)	81 (97.6)	.42
Number of stroke prevention clinic home-based teleconsultations				.74
Once only	52 (47.3)	13 (50)	39 (46.4)	
2-4 times	50 (45.5)	12 (46.2)	38 (45.2)	
≥5 times	8 (7.3)	1 (3.8)	7 (8.3)	
Patient portal use before the COVID-19 pandemic				.71
Never	94 (85.5)	23 (92)	71 (85.5)	
1-2 times	10 (9.1)	2 (8)	8 (9.6)	
3-5 times	3 (2.7)	0 (0)	3 (3.6)	
>5 times	1 (0.9)	0 (0)	1 (1.2)	
Telemedicine use before the COVID-19 pandemic				.35
Never	92 (83.6)	21 (80.8)	71 (86.6)	
1-2 times	7 (6.4)	2 (7.7)	5 (6.1)	
3-5 times	6 (5.5)	1 (3.8)	5 (6.1)	
>5 times	3 (2.7)	2 (7.7)	1 (1.2)	
Previsit contact by the stroke prevention clinic (no)	94 (85.5)	24 (92.3)	70 (85.4)	.36
Owned digital equipment at home (yes)	79 (71.8)	19 (73.1)	60 (72.3)	.94

[a]Note that the percentages are based on denominators that vary from the overall sample size of 110 owing to missing data.

Findings From the Quantitative Analysis

Overall, the instrument was reliable as the Cronbach α reliability analysis of the SERVPERF questionnaire was .894 (Multimedia Appendix 7), which indicated an excellent level of reliability of the instrument [39]. The adjusted R^2 value was 0.76, indicating that the 5 SERVPERF dimensions could explain 76% of the variation in the global satisfaction score. The mean global satisfaction score was 2.5 (SD 0.65) for the low-satisfaction group and 4.40 (SD 0.49) for the high-satisfaction group. To examine the explanatory variables that were significantly associated with the binary outcome variable, consent from the substitute decision maker (SDM), language barrier, age group, survey method, and 5 service dimensions were entered into the final forward stepwise logistic regression model. The adjusted R^2 indicated that 69% of the variance could be explained in the final model. Table 4 illustrates that consent from the SDM (adjusted odds ratio [AOR] 6.59, 95% CI 1.45-29.93; *P*=.01) was positively associated (β=1.89) and the responsiveness (AOR 0.03, 95% CI 0.006-0.188; *P*<.001) and empathy (AOR 0.12, 95% CI 0.02-0.80; *P*=.03) dimensions were negatively associated (β=−3.37 for responsiveness; β=−2.15 for empathy) with dissatisfaction with the home-based teleconsultation service quality (Table 4). The odds of dissatisfaction for participants who consented to the survey through their SDM were 6.59 (95% CI 1.45-29.93) compared with those who consented themselves. Every one-unit increase in the responsiveness dimension score decreased the odds of dissatisfaction by 0.03 (95% CI 0.006-0.19) when other variables were held constant. Every one-unit increase in the empathy dimension score decreased the odds of dissatisfaction by 0.12 (95% CI 0.02-0.8) when other variables were held constant.

Table 4. The forward stepwise binary logistic regression model.

Variable	β (SE)	*P* value	AOR[a] (95% CI)
Consent from SDM[b]	1.89 (0.772)	.01[c]	6.59 (1.45-29.93)
Responsiveness[d]	−3.37 (0.867)	<.001[e]	0.03 (0.006-0.19)
Empathy[f]	−2.15 (0.986)	.03[c]	0.12 (0.02-0.80)

[a]AOR: adjusted odds ratio.

[b]SDM: substitute decision maker. Consent from SDM indicates that the survey was answered with the help of an SDM.

[c]*P*<.05.

[d]Responsiveness is a service performance model dimension regarding the willingness to help customers and provide a prompt service.

[e]*P*<.001.

[f]Empathy is a service performance model dimension regarding providing individual care and attention to customers.

The significant characteristics of the participants who had their SDM sign the consent form and help them answer the survey are listed in Table 5. The participants whose SDM provided consent to help them answer the survey were more likely to answer a web-based survey (χ^2_2=15.6; *P*<.001) and have a language barrier (χ^2_2=15.6; *P*<.001), hearing impairment (χ^2_2=3.9; *P*=.048), or hearing that affected telephone conversations (χ^2_2=5.6; *P*<.02) and were less likely to drive (χ^2_2=7.0; *P*=.04). Tangibles was the only statistically significant SERVPERF dimension (*P*<.001). The participants who consented through their SDM had a shorter travel distance, were older, and were more likely to have residual stroke symptoms.

Table 5. Comparison of variables between surveys for which consent was obtained from the substitute decision maker (SDM) and surveys for which consent was obtained from the patient themselves (difference of >15%).

Characteristic	Consent from SDM[a] (n=53)	Consent from patient (n=57)	*P* value
Web-based survey, n (%)	50 (94)	36 (63)	<.001[b]
Age (y), mean (SD)	73.75 (13.085)	70.60 (10.61)	.17
Male sex, n (%)	34 (64)	27 (47)	.08
Distance (km), mean (SD)	15.897 (16.525)	19.435 (16.48)	.26
Driving, n (%)	27 (55)[c]	40 (74)[d]	.04[e]
Language barrier, n (%)	10 (19)	0 (0)	<.001[b]
Hearing impaired, n (%)	18 (34)	10 (18)	.048
Hearing impairment affecting phone conversations, n (%)	12 (67)[f]	2 (20)[g]	.02[e]
Residual stroke symptoms, n (%)	17 (49)[h]	10 (28)[i]	.07
Tangibles, mean (SD)	3.21 (0.83)	3.67 (0.61)	<.001[b]
Reliability, mean (SD)	3.87 (0.82)	3.94 (0.56)	.68
Responsiveness, mean (SD)	3.76 (0.94)	3.76 (0.59)	.98
Assurance, mean (SD)	3.99 (0.78)	4.04 (0.53)	.68
Empathy, mean (SD)	3.75 (0.71)	3.96 (0.56)	.34

[a]Consent from SDM indicates that the survey was answered with the help of an SDM.
[b]*P*<.001.
[c]N=49.
[d]N=54.
[e]*P*<.05.
[f]N=18.
[g]N=10.
[h]N=35.
[i]N=36.

Findings From the Content Analysis

Overview

A total of 88.2% (97/110) of patients completed the open-ended questions in the survey, with 25% (24/97) in the low-satisfaction group and 75% (73/97) in the high-satisfaction group. Multimedia Appendices 8 and 9 list the dimensions and subcategories of positive and negative comments among patients with low and high satisfaction scores, respectively. Interestingly, the low- and high-satisfaction groups shared the same dissatisfied service dimensions (assurance, reliability, and empathy) and subcategories. Overall, missing clinical components, inadequate communication, administrative issues, and absence of personal connection were the significant concerns that affected patients' perceived home-based teleconsultation quality at the stroke prevention clinics.

In contrast, the most satisfying service dimensions were assurance, empathy, and responsiveness among the high-satisfaction group. Overall, a competent clinician with effective communication skills and great empathy for patients is crucial for patient-perceived high-quality care in home-based teleconsultation. In addition, convenience, appropriateness to the patient's situation, and timely consultation were important for high satisfaction with home-based teleconsultation.

Future Use of Home-Based Teleconsultation

Assessment of patient preference for future use of home-based teleconsultation under normal circumstances (after the COVID-19 pandemic) showed that 35% (33/95) of participants preferred not to use home-based teleconsultation. Most participants (18/23, 78%) in the low-satisfaction group preferred not to use home-based teleconsultation. In contrast, 56% (40/72) of participants in the high-satisfaction group were willing to use it, and 24% (17/72) indicated that they might use home-based teleconsultation for specific reasons. However, a minority of participants in the high-satisfaction group (15/72, 21%) still refused to use it in normal circumstances, with primary concerns of communication issues, lack of personal connection, and the belief in the superiority of in-person consultations.

Discussion

Principal Findings

Home-based teleconsultation as a form of telemedicine rapidly expanded in many health sectors in Canada during the

COVID-19 pandemic owing to lockdowns and social distancing restrictions [31]. Since the COVID-19 pandemic, home-based teleconsultation has become essential in outpatient service delivery [40]. By April 2020, 77% of Ontario ambulatory visits were conducted using teleconsultation, a total of 77% of ambulatory visits were conducted using a virtual modality [41]. Nearly all (32/33, 97%) Ontario stroke prevention clinics that responded to a province-wide survey from June 2021 to July 2021 (response rate of 33/41, 80%) reported that they had adopted home-based teleconsultation as a service delivery mode in addition to in-person visits since the COVID-19 pandemic [42]. Patient satisfaction should be the priority in future virtual care development and adoption [43]. To our knowledge, this is the first study to use a service quality theoretical lens to assess patient satisfaction with home-based teleconsultation in outpatient stroke care during the pandemic. Many studies of patient surveys during the COVID-19 pandemic have found that most patients and clinicians reported positive experiences with teleconsultation at outpatient neurology services during the COVID-19 pandemic [31]. However, no study has investigated outpatient stroke prevention services or examined the service quality of home-based teleconsultation during the COVID-19 pandemic. The patient population of stroke prevention clinics and the disease characteristics differ from those of other chronic neurological diseases. For instance, patients at stroke prevention clinics have a unique mix of acuity (such as early identification of large vessel occlusion and cardiac source of embolism) and chronic disease management (eg, hypertension, dyslipidemia, diabetes, and lifestyle management), and most are older adults. Owing to health resource disparity, timely access to outpatient magnetic resonance imaging is not always feasible for minor strokes. When referred to stroke prevention clinics, patients with TIA have transient neurological symptoms, unremarkable brain images, and normal physical examinations. History taking is essential in patients with TIA. The unique patient population and characteristics may pose different challenges in home-based teleconsultation, especially for newly referred patients. Our study found that the participants who were older (mean age 72.12, SD 11.92 years) and mostly newly referred (101/110, 91.8%) and used the telephone modality (107/109, 98.2%) were satisfied with the home-based teleconsultation provided by the stroke prevention clinics during the COVID-19 pandemic. We identified patient-reported factors that affected their satisfaction with the service quality of home-based teleconsultation. Our study filled these research gaps.

Responsiveness was the most statically significant dimension in our quantitative results and is an influential factor for a positive experience. Convenience was the main subtheme of the responsiveness dimension in the high-satisfaction group. The patients in the low-satisfaction group tended to live closer to the stroke prevention clinics, with an average distance of 13.53 (SD 11.57) km, than those in the high-satisfaction group (mean 19.03, SD 17.61 km). Even though they were less likely to drive and had some communication barriers (more likely to have language barriers or hearing impairments), convenience was not a positive factor influencing their satisfaction. Our content analysis supported that convenience, by saving time, travel, and energy, influenced patients' positive perceptions of the personal benefits of home-based teleconsultation during the pandemic [31]. Our findings were consistent with those of studies conducted before the COVID-19 pandemic [44]. A systematic review of digital experience also found that convenience is one of the motivating factors contributing to a positive digital patient experience [45]. The convenience of home-based teleconsultation is an influential factor in swaying patients' satisfaction with service quality to the positive side and their preference for teleconsultation [31]. However, convenience is not equivalent to good quality of care. We need to consider the effect of convenience when assessing patients' preferences and evaluating patient satisfaction with the service quality of home-based teleconsultation.

Second, the empathy dimension was a significant factor in both the quantitative and qualitative analyses. Dissatisfaction feedback in the empathy dimension was found in both the low- and high-satisfaction groups. Our study participants used mostly the telephone modality, where the lack of nonverbal cues may be associated with a profound concern about the lack of personal connection, which is the primary subcategory of the empathy dimension. As indicated by existing studies, the replacement of interpersonal connection and a lack of physical human contact are negatively associated with digital patient experiences [45]. The literature shows that empathy can have powerful effects on positive patient outcomes and satisfaction [46,47]. There is a concern that the digitalization of health care services could primarily lead to a decrease in the expression of empathy [46]. A study on patient satisfaction with tele-obstetric care found that a desire for personal connection via face-to-face interaction with a clinician was a critical motivation for selecting in-person versus teleconsultation care modalities [48]. Our findings are in line with those of the previous literature. The lack of personal connection in our content analysis could explain the negative relationship between the empathy dimension and dissatisfaction. An interesting finding from our content analysis was that even 98% (81/83) of the participants in the high-satisfaction group used the telephone modality; they expressed overwhelmingly more positive than negative comments (34 vs 12) on the empathy dimension. The clinician's empathy skills may significantly enhance patient experiences even with a low-technology modality. Empathy skill training for clinicians, primarily through computer-mediated communications, is a key area to study in the future [46].

Third, although the assurance dimension was not found to be statistically significant between the low- and high-satisfaction groups, this was likely due to no real difference between the 2 groups. Our content analysis indicated that the assurance dimension was one of the most important SERVPERF dimensions in both the low- and high-satisfaction groups. The subcategories raised from the content analysis revealed that the patients' concerns in the assurance dimension were the missing clinical components—especially physical examination—and inadequate communication. The view of the inferior quality of a remote examination among clinicians was dominant in outpatient neurology teleconsultation before the COVID-19 pandemic [49]. This is likely why most teleconsultations were performed for follow-up patients with chronic neurological diseases before the COVID-19 pandemic [50]. Compared with

before the COVID-19 pandemic, home-based teleconsultation has been widely used in both new and follow-up patients at home without the luxury of having a health care professional assisting a teleconsultation since the COVID-19 pandemic [31]. The lack of a physical examination and inadequate communication could impede the clinician's ability to diagnose and formulate a treatment plan, especially for new patients [31]. In addition, the use of video is more challenging than the use of a telephone because of the rapid adaptation and lack of preparation. According to the Ontario stroke prevention clinic web survey, nearly half of the clinics use only the telephone [42]. Video consultations enable a certain degree of remote examination and may facilitate better communication, whereas telephone-only visits limit clinical assessment and communication. Telephone consultations lack body language and physical prompts, which could negatively affect the communication between the clinician and patient. However, most of the participants in our study used telephone consultations (107/110, 97.3%), and their overall satisfaction was high (3.9/5). Future studies could consider patient satisfaction when using a telephone-only modality for new referrals in this patient population under normal circumstances.

The appropriateness of patient selection is a critical factor in high-quality home-based teleconsultation from patients' perspectives. Our statistical findings indicated that patients who required help from their SDM to consent and answer the survey were positively associated with dissatisfaction. The SDM may have chosen a web survey, as the participants had difficulty answering a telephone survey because of language barriers, hearing impairments, or communication difficulties from residual stroke symptoms (such as aphasia, apraxia, or mild cognitive impairment; Table 5). Moreover, the patients who needed their SDM to consent and help them answer the survey scored significantly lower in the tangibles dimension, which may indicate that the participants had low comfort levels with technology and technical difficulties even with the telephone modality. The consent from an SDM has many unknown characteristics and requires further exploration in future research.

Similarly, the analysis showed that there were no statistically significant differences in the reliability dimension between the low- and high-satisfaction groups. However, it is important to note that issues primarily related to administration, which fall under the reliability dimension, had a negative impact on both the low- and high-satisfaction groups. This possibility aligns with findings from a scoping review that the lack of proper administrative support has harmed clinicians' teleconsultation satisfaction [31]. Our findings indicate that it also negatively affects patient satisfaction. Some of the low satisfaction may be due to the abrupt change to teleconsultation because of the COVID-19 pandemic and the lack of clinical and patient preparation. We know that some administrative problems are not unique to home-based teleconsultation, as they occur during in-person visits. Home-based teleconsultation may have increased the workload by keeping pace with the transitioning workflow among telephone, video, and in-person visits, contributing to the maladaptation of home-based teleconsultation [51]. Establishing a new care pathway for home-based teleconsultation may streamline the administrative workflow.

The assurance, empathy, and reliability dimensions all had the most negative comments in both the low- and high-satisfaction groups, and the high-satisfaction group had the most positive remarks in the assurance and empathy dimensions in the content analysis, which showed a double-edged effect. This finding might indicate that different key subcategories have buffer effects on improving patient satisfaction with the service quality of home-based teleconsultation. This may reflect the advantage of using both quantitative and qualitative data to provide diverse types of information [37]. By comparing them side by side, the qualitative analysis may provide insights to explain the quantitative findings.

The Future of Stroke Prevention Clinics' Service Delivery Mode

A combination of teleconsultation and in-person visits for outpatient stroke prevention care is the future. Our study showed that 45% (43/95) of participants were willing to use and 18% (17/95) would consider using home-based teleconsultation in future nonpandemic conditions. A study examining patient preference for telehealth for nonemergent health issues after the COVID-19 pandemic concluded that patients were generally willing to use video but preferred in-person visits [52]. A patient-centered service should be delivered by offering the patient a choice [20]. Virtual care provides an opportunity to design a health system that is actually patient-centered [43]. A combination of in-person visits and home-based teleconsultation—a hybrid care model—could best meet patient needs by improving efficiency and capacity without added risk [53,54].

Hybrid care should be sustainable in practice settings to ensure patient care quality, equity, and justice [53]. To avoid increasing the digital divide, a telephone may be favorable instead of video calls for older patients and those with a lower education or income and from racial and ethnic minority groups [55]. However, we should refrain from creating a 2-tiered health care system in which high-income individuals receive video consultations and low-income individuals receive phone consultations. Patients should receive the right care in the right setting, at the right time, and with the right mode; the cost of the service should be reduced; and the best clinical practice guidelines should be followed [56]. Hybrid care could be a balanced approach to achieving a high-performance health care system. Patients can choose the best model by considering flexible options, and clinicians can offer individualized recommendations for optimal care modalities [54].

Limitations

Our study has several limitations. First, the cross-sectional survey only provides a snapshot of a phenomenon and cannot determine the temporal relationship between the dependent and independent variables. Second, the participants in this study may not be generalizable to patients of other stroke prevention clinics in Ontario, notably in areas with different health resources such as urban versus very remote rural centers. In addition, the open-ended responses to the web survey were very

brief, limiting our ability to gain a deeper understanding of their experiences. Next, as we surveyed patients who had had a home-based teleconsultation within 6 months, recall bias is possible [57]. Only patients who had a home-based teleconsultation in January 2021 and February 2021 received the survey 4 to 5 months later; the following patients received their survey 2 and a half months after the home-based teleconsultation on average (range 1-3 months). In addition, there is potential nonresponse bias, as web surveys usually have a low response rate [57]. A meta-analysis comparing web survey response rates concluded that the average response rate for web surveys was approximately 11% [58]. Our web survey had a 35.9% (104/290) response rate, and the telephone survey yielded a 44% (24/54) response rate.

Overall, although our study may only reflect part of the concept of satisfaction with service quality because of its complexity, it provided substantial insights into areas for quality improvements from patients' point of view. The literature suggests that patient satisfaction is a multidimensional concept that still needs to be fully defined. The patient satisfaction scores may reflect the demographic mix and clinical and psychological picture of the patients served by a medical service [59]. Our study attempted to use a theory-guided quantitative and qualitative analysis to reveal the relationship and explanation of such a complex phenomenon. Despite the problems of using patient satisfaction to assess service quality, its measurement provides unique information regarding the care process as seen through the patients' eyes [59]. Patients still provide the best source of accurate information on the care they receive [60]. In the absence of choice in public-funded health care, our survey gave patients a voice to indicate their preferences [21].

Conclusions

Our findings highlighted 2 crucial service quality dimensions (responsiveness and empathy) that were negatively statistically significantly associated with patient dissatisfaction. Moreover, we identified that a survey consented to by an SDM was positively associated with dissatisfaction. In addition, there were 4 subcategories related to patient dissatisfaction (missing clinical activities, inadequate communication, administrative process issues, and absence of personal connection). We anticipate that appropriate patient selection, consideration of patient preferences, a streamlined home-based teleconsultation administrative workflow, and a competent clinician with communication and empathy skills are essential for achieving high satisfaction with home-based teleconsultation. These factors could be considered when designing home-based teleconsultation services to enhance patient experiences of stroke prevention care.

Acknowledgments

This work was supported by 2 stroke prevention clinics. The authors acknowledge especially Dr Courtney Scott, a neurologist, for helping with patient recruitment and Lorri Reynold, the director of Patient Experience, Professional Practice, and Research, for arranging personnel for patient recruitment. This work was also supported by the University of Waterloo, especially by Gaya Bin Noon, a master's student at the School of Public Health Sciences of the University of Waterloo, who served as the research assistant for this study.

Data Availability

Data are available upon request from the authors.

Conflicts of Interest

GM is a clinical practitioner in one of the stroke prevention clinics but was not involved in patient contact during the research, including the recruitment and telephone interview processes.

Multimedia Appendix 1
Survey consent sheet for participants.
[DOCX File , 26 KB - cardio_v8i1e51439_app1.docx]

Multimedia Appendix 2
Cover letter for mail-in package.
[DOCX File , 19 KB - cardio_v8i1e51439_app2.docx]

Multimedia Appendix 3
The description of 18 items used in the service performance model questionnaire.
[DOCX File , 13 KB - cardio_v8i1e51439_app3.docx]

Multimedia Appendix 4
Survey: the patient's perception of teleconsultation service quality at stroke prevention clinic during COVID-19.
[DOCX File , 24 KB - cardio_v8i1e51439_app4.docx]

Multimedia Appendix 5
Six questions to check the clarity of the survey questions.
[DOCX File , 13 KB - cardio_v8i1e51439_app5.docx]

Multimedia Appendix 6
Definitions of 10 demographic, 7 clinical, and 6 technical-related independent variables.
[DOCX File , 14 KB - cardio_v8i1e51439_app6.docx]

Multimedia Appendix 7
Cronbach α reliability statistics and item-total statistics of the service performance model questionnaire.
[DOCX File , 14 KB - cardio_v8i1e51439_app7.docx]

Multimedia Appendix 8
Positive and negative categories among patients with low global satisfaction.
[DOCX File , 17 KB - cardio_v8i1e51439_app8.docx]

Multimedia Appendix 9
Positive and negative categories among patients with high global satisfaction.
[DOCX File , 18 KB - cardio_v8i1e51439_app9.docx]

References

1. GBD 2019 Diseases and Injuries 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]
2. Leading causes of death, total population, by age group. Statistics Canada. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310039401 [accessed 2023-06-06]
3. GBD 2016 Stroke Collaborators. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 May;18(5):439-458 [FREE Full text] [doi: 10.1016/S1474-4422(19)30034-1] [Medline: 30871944]
4. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. Lancet 2010 Jul 10;376(9735):112-123 [FREE Full text] [doi: 10.1016/s0140-6736(10)60834-3]
5. Hill MD, Yiannakoulias N, Jeerakathil T, Tu JV, Svenson LW, Schopflocher DP. The high risk of stroke immediately after transient ischemic attack: a population-based study. Neurology 2004 Jun 08;62(11):2015-2020. [doi: 10.1212/01.wnl.0000129482.70315.2f] [Medline: 15184607]
6. CorHealth Ontario. Ontario stroke report FY 2019-20. Stroke Network of Southeastern Ontario. 2021 Jun. URL: https://www.strokenetworkseo.ca/sites/default/files/files/ontario_stroke_report_2019-20_-_final.pdf [accessed 2023-07-12]
7. Core elements – Ontario stroke prevention clinics. CorHealth. URL: https://www.corhealthontario.ca/resources-for-healthcare-planners-&-providers/piwp/Ontario-SPC-Core-Elements(2018).pdf [accessed 2022-03-11]
8. Webster F, Saposnik G, Kapral MK, Fang J, O'Callaghan C, Hachinski V. Organized outpatient care: stroke prevention clinic referrals are associated with reduced mortality after transient ischemic attack and ischemic stroke. Stroke 2011 Nov;42(11):3176-3182. [doi: 10.1161/STROKEAHA.111.621524] [Medline: 21921281]
9. Hall R, Khan F, O'Callaghan C, Kapral MK, Hodwitz K, Kapila S, et al. Ontario stroke evaluation report 2013: spotlight on secondary stroke prevention and care. Institute for Clinical Evaluative Sciences. 2013. URL: https://ontariostrokenetwork.ca/wp-content/uploads/2014/06/Full-report-2013.pdf [accessed 2023-03-12]
10. Appireddy R, Khan S, Leaver C, Martin C, Jin A, Durafourt BA, et al. Home virtual visits for outpatient follow-up stroke care: cross-sectional study. J Med Internet Res 2019 Oct 07;21(10):e13734 [FREE Full text] [doi: 10.2196/13734] [Medline: 31593536]
11. Teleconsultations during a pandemic. Pan American Health Organization. URL: https://iris.paho.org/bitstream/handle/10665.2/52006/covid-19-teleconsultations-eng.pdf?sequence=4 [accessed 2023-07-17]
12. Virtual care in Canada: progress and potential. Canadian Medical Association. 2022 Feb. URL: https://www.cma.ca/sites/default/files/2022-02/Virtual-Care-in-Canada-Progress-and-Potential-EN.pdf [accessed 2023-01-16]
13. Different strokes: recovery triumphs and challenges at any age. Heart and Stroke Foundation of Canada. 2017. URL: https://www.heartandstroke.ca/-/media/pdf-files/canada/stroke-report/strokereport2017en.ashx[accessed [accessed 2023-01-21]
14. COVID-19 baseline survey: who participated? CLSA Dashboard. URL: https://dashboard.clsa-elcv.ca/clsa-dashboard-en/ [accessed 2024-01-09]
15. Health service delivery. World Health Organization. URL: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/health-service-delivery [accessed 2023-03-07]

16. Mosadeghrad AM. A conceptual framework for quality of care. Mater Sociomed 2012;24(4):251-261 [FREE Full text] [doi: 10.5455/msm.2012.24.251-261] [Medline: 23922534]
17. Zeithaml VA, Bitner MJ, Gremler DD. Services Marketing: Integrating Customer Focus Across the Firm. New York, NY: McGraw-Hill Education; 2017.
18. Grönroos C. A service quality model and its marketing implications. Eur J Mark 1984 Apr 1;18(4):36-44. [doi: 10.1108/EUM0000000004784]
19. van Oppen JD, Valderas JM, Mackintosh NJ, Conroy SP. Patient-reported outcome and experience measures in geriatric emergency medicine. Z Gerontol Geriatr 2021 Mar;54(2):122-124 [FREE Full text] [doi: 10.1007/s00391-020-01777-4] [Medline: 32939573]
20. Ferreira PL, Raposo V, Tavares AI. Primary health care patient satisfaction: explanatory factors and geographic characteristics. Int J Qual Health Care 2020 Apr 27;32(2):93-98 [FREE Full text] [doi: 10.1093/intqhc/mzz134] [Medline: 32047931]
21. Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L, et al. The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. Health Technol Assess 2002;6(32):1-244 [FREE Full text] [doi: 10.3310/hta6320] [Medline: 12925269]
22. Parasuraman A, Zeithaml VA, Berry LL. SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. J Retail 1988;64(1):12-40.
23. Cronin JJ, Taylor SA. Measuring service quality: a reexamination and extension. J Mark 2018 Nov 28;56(3):55-68 [FREE Full text] [doi: 10.1177/002224299205600304]
24. Deming WE. Improvement of quality and productivity through action by management. Natl Prod Rev 1981;1(1):12-22 [FREE Full text] [doi: 10.1002/npr.4040010105]
25. Camilleri D, O'Callaghan M. Comparing public and private hospital care service quality. Int J Health Care Qual Assur Inc Leadersh Health Serv 1998 Jul 01;11(4-5):127-133 [FREE Full text] [doi: 10.1108/09526869810216052] [Medline: 10185325]
26. Aagja JP, Garg R. Measuring perceived service quality for public hospitals (PubHosQual) in the Indian context. Int J Pharm Healthc Mark 2010 Apr 06;4(1):60-83 [FREE Full text] [doi: 10.1108/17506121011036033]
27. Itumalla R, Acharyulu GV, Shekhar BR. Development of HospitalQual: a service quality scale for measuring inpatient services in hospital. OSCM 2014;7(2):54-63 [FREE Full text] [doi: 10.31387/oscm0170108]
28. Brady MK, Cronin JJJ, Brand RR. Performance-only measurement of service quality: a replication and extension. J Bus Res 2002 Jan;55(1):17-31 [FREE Full text] [doi: 10.1016/s0148-2963(00)00171-5]
29. Zeithaml VA, Parasuraman A, Berry LL. Delivering Quality Service: Balancing Customer Perceptions and Expectations. New York, NY: Free Press; 1990.
30. Yin SY, Huang KK, Shieh JI, Liu YH, Wu HH. Telehealth services evaluation: a combination of SERVQUAL model and importance-performance analysis. Qual Quant 2015 Feb 8;50(2):751-766 [FREE Full text] [doi: 10.1007/s11135-015-0174-4]
31. Meng G, McAiney C, Perlman CM, McKillop I, Tisseverasinghe T, Chen HH. Service process factors affecting patients' and clinicians' experiences on rapid teleconsultation implementation in out-patient neurology services during COVID-19 pandemic: a scoping review. BMC Health Serv Res 2022 Apr 22;22(1):534 [FREE Full text] [doi: 10.1186/s12913-022-07908-4] [Medline: 35459134]
32. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
33. IBM SPSS Statistics for Windows, version 28.0. IBM Corp. 2021. URL: https://www.ibm.com/products/spss-statistics [accessed 2023-12-20]
34. Khan JA, Van Aelst S, Zamar RH. Building a robust linear model with forward selection and stepwise procedures. Comput Stat Data Anal 2007 Sep 15;52(1):239-248 [FREE Full text] [doi: 10.1016/j.csda.2007.01.007]
35. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288 [FREE Full text] [doi: 10.1177/1049732305276687] [Medline: 16204405]
36. Potter WJ, Levine‐Donnerstein D. Rethinking validity and reliability in content analysis. J Appl Commun Res 2009 May 21;27(3):258-284. [doi: 10.1080/00909889909365539]
37. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: SAGE Publications; 2017.
38. Downey RG, King C. Missing data in Likert ratings: a comparison of replacement methods. J Gen Psychol 1998 Apr 30;125(2):175-191 [FREE Full text] [doi: 10.1080/00221309809595542] [Medline: 9935342]
39. Cronbach LJ, Shapiro K. Designing Evaluations of Educational and Social Programs. Hoboken, NJ: Jossey-Bass; 1982.
40. Fonseca E, Quintana M, Lallana S, Luis Restrepo J, Abraira L, Santamarina E, et al. Epilepsy in time of COVID-19: a survey-based study. Acta Neurol Scand 2020 Dec;142(6):545-554 [FREE Full text] [doi: 10.1111/ane.13335] [Medline: 32799337]
41. Mehrotra A, Bhatia RS, Snoswell CL. Paying for Telemedicine After the Pandemic. JAMA 2021 Feb 02;325(5):431-432 [FREE Full text] [doi: 10.1001/jama.2020.25706] [Medline: 33528545]

42. Provincial SPC Working Group. Virtual care in secondary stroke prevention: considerations and recommendations for hybrid models of care. CorHealth Ontario. 2022. URL: https://www.corhealthontario.ca/Virtual-Care-in-SPC-Considerations-and-Recommendations-2022.pdf [accessed 2023-10-23]
43. Enhancing equitable access to virtual care in Canada: principle-based recommendations for equity. Health Canada. 2021 Jun 29. URL: https://www.canada.ca/content/dam/hc-sc/documents/corporate/transparency_229055456/health-agreements/bilateral-agreement-pan-canadian-virtual-care-priorities-covid-19/template-ett-report-docx-eng.pdf [accessed 2023-01-16]
44. Fleischhacker CL. Patient satisfaction with telehealth services compared to in-office visits: a systematic literature review. Minnesota State University. 2020. URL: https://cornerstone.lib.mnsu.edu/etds/982/ [accessed 2023-07-21]
45. Wang T, Giunti G, Melles M, Goossens R. Digital patient experience: umbrella systematic review. J Med Internet Res 2022 Aug 04;24(8):e37952 [FREE Full text] [doi: 10.2196/37952] [Medline: 35925651]
46. Terry C, Cain J. The emerging issue of digital empathy. Am J Pharm Educ 2016 May 25;80(4):58 [FREE Full text] [doi: 10.5688/ajpe80458] [Medline: 27293225]
47. Pealing L, Tempest HV, Howick J, Dambha-Miller H. Technology: a help or hindrance to empathic healthcare? J R Soc Med 2018 Nov;111(11):390-393 [FREE Full text] [doi: 10.1177/0141076818790669] [Medline: 30175938]
48. Pflugeisen BM, Mou J. Patient satisfaction with virtual obstetric care. Matern Child Health J 2017 Jul 7;21(7):1544-1551 [FREE Full text] [doi: 10.1007/s10995-017-2284-1] [Medline: 28176034]
49. Dorsey ER, Glidden AM, Holloway MR, Birbeck GL, Schwamm LH. Teleneurology and mobile technologies: the future of neurological care. Nat Rev Neurol 2018 May;14(5):285-297 [FREE Full text] [doi: 10.1038/nrneurol.2018.31] [Medline: 29623949]
50. Patel UK, Malik P, DeMasi M, Lunagariya A, Jani VB. Multidisciplinary approach and outcomes of tele-neurology: a review. Cureus 2019 Apr 08;11(4):e4410 [FREE Full text] [doi: 10.7759/cureus.4410] [Medline: 31205830]
51. Kummer BR, Sweetnam C, Vickrey BG, Naasan G, Harvey D, Gallagher K, et al. Teleneurology expansion in response to the COVID-19 outbreak at a tertiary health system in New York City. Neur Clin Pract 2021 Apr 09;11(2):e102-e111. [doi: 10.1212/cpj.0000000000001057]
52. Predmore ZS, Roth E, Breslau J, Fischer SH, Uscher-Pines L. Assessment of patient preferences for telehealth in post-COVID-19 pandemic health care. JAMA Netw Open 2021 Dec 01;4(12):e2136405 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.36405] [Medline: 34851400]
53. Kim J, Webb N, Matos N, Furuya R. Learning from the COVID-19 pandemic: designing and implementing a telehealth-in person hybrid care model. Free Clin Res Collect 2020;6(1).
54. Westwood AR. Is hybrid telehealth model the next step for private healthcare in India? Health Serv Insights 2021 Sep 01;14:11786329211043301 [FREE Full text] [doi: 10.1177/11786329211043301] [Medline: 34483663]
55. Danila MI, Sun D, Jackson LE, Cutter G, Jackson EA, Ford EW, et al. Satisfaction with modes of telemedicine delivery during COVID-19: a randomized, single-blind, parallel group, noninferiority trial. Am J Med Sci 2022 Nov;364(5):538-546 [FREE Full text] [doi: 10.1016/j.amjms.2022.06.021] [Medline: 35793732]
56. Better health, better care, better value for all: refocusing health care reform in Canada. Health Council of Canada. 2013 Sep. URL: https://neltoolkit.rnao.ca/sites/default/files/Better%20Health_Better%20Care_Better%20Value%20For%20All_Refocusing%20Health%20Care%20in%20Canada%20September%202013.pdf [accessed 2023-05-03]
57. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. Nephron Clin Pract 2010 Apr 21;115(2):c94-c99 [FREE Full text] [doi: 10.1159/000312871] [Medline: 20407272]
58. Manfreda KL, Bosnjak M, Berzelak J, Haas I, Vehovar V. Web surveys versus other survey modes: a meta-analysis comparing response rates. Int J Mark Res 2008 Jan 01;50(1):79-104 [FREE Full text] [doi: 10.1177/147078530805000107]
59. Schoenfelder T. Patient satisfaction: a valid indicator for the quality of primary care? Primary Health Care 2012;02(04):1000e106. [doi: 10.4172/2167-1079.1000e106]
60. Epstein KR, Laine C, Farber NJ, Nelson EC, Davidoff F. Patients' perceptions of office medical practice: judging quality through the patients' eyes. Am J Med Qual 2016 Jun 30;11(2):73-80 [FREE Full text] [doi: 10.1177/0885713x9601100204]

Abbreviations

AOR: adjusted odds ratio
REDCap: Research Electronic Data Capture
SDM: substitute decision maker
SERVPERF: service performance model
SERVQUAL: service quality model
TIA: transient ischemic attack

Edited by A Mavragani; submitted 04.08.23; peer-reviewed by Z Wang, E van der Velde; comments to author 30.11.23; revised version received 21.12.23; accepted 03.01.24; published 16.02.24.

Please cite as:
Meng G, McAiney C, McKillop I, Perlman CM, Tsao SF, Chen H
Factors That Influence Patient Satisfaction With the Service Quality of Home-Based Teleconsultation During the COVID-19 Pandemic: Cross-Sectional Survey Study
JMIR Cardio 2024;8:e51439
URL: https://cardio.jmir.org/2024/1/e51439
doi:10.2196/51439
PMID:38363590

Original Paper

A Multidisciplinary Assessment of ChatGPT's Knowledge of Amyloidosis: Observational Study

Ryan C King[1], MD; Jamil S Samaan[2], MD; Yee Hui Yeo[2], MSc, MD; Yuxin Peng[3], BS; David C Kunkel[4], MD; Ali A Habib[5], MD; Roxana Ghashghaei[1], MD

[1]Division of Cardiology, Department of Medicine, University of California, Irvine Medical Center, Orange, CA, United States

[2]Karsh Division of Gastroenterology and Hepatology, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, United States

[3]School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

[4]GI Motility and Physiology Program, Division of Gastroenterology, University of California, San Diego, La Jolla, CA, United States

[5]Division of Neurology, University of California, Irvine Medical Center, Orange, CA, United States

Corresponding Author:
Ryan C King, MD
Division of Cardiology
Department of Medicine
University of California, Irvine Medical Center
101 The City Drive South
Orange, CA, 92868
United States
Phone: 1 714 456 7890
Email: kingrc@hs.uci.edu

Abstract

Background: Amyloidosis, a rare multisystem condition, often requires complex, multidisciplinary care. Its low prevalence underscores the importance of efforts to ensure the availability of high-quality patient education materials for better outcomes. ChatGPT (OpenAI) is a large language model powered by artificial intelligence that offers a potential avenue for disseminating accurate, reliable, and accessible educational resources for both patients and providers. Its user-friendly interface, engaging conversational responses, and the capability for users to ask follow-up questions make it a promising future tool in delivering accurate and tailored information to patients.

Objective: We performed a multidisciplinary assessment of the accuracy, reproducibility, and readability of ChatGPT in answering questions related to amyloidosis.

Methods: In total, 98 amyloidosis questions related to cardiology, gastroenterology, and neurology were curated from medical societies, institutions, and amyloidosis Facebook support groups and inputted into ChatGPT-3.5 and ChatGPT-4. Cardiology- and gastroenterology-related responses were independently graded by a board-certified cardiologist and gastroenterologist, respectively, who specialize in amyloidosis. These 2 reviewers (RG and DCK) also graded general questions for which disagreements were resolved with discussion. Neurology-related responses were graded by a board-certified neurologist (AAH) who specializes in amyloidosis. Reviewers used the following grading scale: (1) comprehensive, (2) correct but inadequate, (3) some correct and some incorrect, and (4) completely incorrect. Questions were stratified by categories for further analysis. Reproducibility was assessed by inputting each question twice into each model. The readability of ChatGPT-4 responses was also evaluated using the *Textstat* library in Python (Python Software Foundation) and the *Textstat readability* package in R software (R Foundation for Statistical Computing).

Results: ChatGPT-4 (n=98) provided 93 (95%) responses with accurate information, and 82 (84%) were comprehensive. ChatGPT-3.5 (n=83) provided 74 (89%) responses with accurate information, and 66 (79%) were comprehensive. When examined by question category, ChatGTP-4 and ChatGPT-3.5 provided 53 (95%) and 48 (86%) comprehensive responses, respectively, to "general questions" (n=56). When examined by subject, ChatGPT-4 and ChatGPT-3.5 performed best in response to cardiology questions (n=12) with both models producing 10 (83%) comprehensive responses. For gastroenterology (n=15), ChatGPT-4 received comprehensive grades for 9 (60%) responses, and ChatGPT-3.5 provided 8 (53%) responses. Overall, 96 of 98 (98%) responses for ChatGPT-4 and 73 of 83 (88%) for ChatGPT-3.5 were reproducible. The readability of ChatGPT-4's responses ranged from 10th to beyond graduate US grade levels with an average of 15.5 (SD 1.9).

Conclusions: Large language models are a promising tool for accurate and reliable health information for patients living with amyloidosis. However, ChatGPT's responses exceeded the American Medical Association's recommended fifth- to sixth-grade reading level. Future studies focusing on improving response accuracy and readability are warranted. Prior to widespread implementation, the technology's limitations and ethical implications must be further explored to ensure patient safety and equitable implementation.

(JMIR Cardio 2024;8:e53421) doi:10.2196/53421

KEYWORDS

amyloidosis; ChatGPT; large language models; cardiology; gastroenterology; neurology; artificial intelligence; multidisciplinary care; assessment; patient education; large language model; accuracy; reliability; accessibility; educational resources; dissemination; gastroenterologist; cardiologist; medical society; institution; institutions; Facebook; neurologist; reproducibility; amyloidosis-related

Introduction

Background

Amyloidosis is a rare, multisystem disease that comprises several subtypes including secondary amyloidosis, light chain amyloidosis, and ATTR (transthyretin amyloidosis), with the latter 2 being the most common but often underdiagnosed [1]. Light chain amyloidosis is diagnosed in 2500 to 5000 individuals annually in the United States, while the exact incidence of ATTR and secondary amyloidosis remains unknown due to challenges and delays in diagnosis stemming from a broad range of symptoms affecting multiple organ systems [2,3]. Diagnosing and caring for patients living with amyloidosis necessitate effective multidisciplinary collaboration between specialists in fields including but not limited to cardiology, gastroenterology, and neurology [4].

Due to amyloidosis being a rare disease, patients may be at risk for decreased health literacy regarding their condition. A notable scarcity of patient education materials (PEMs) exists for rare diseases compared to common ones, with one study showing nearly a 10-fold difference in the availability of PEMs related to rare diseases, which has been shown to adversely affect health outcomes [5]. According to the Centers for Disease Control and Prevention [6], improved health literacy could prevent up to 1 million hospitalizations annually and save US $25 billion in total health care costs.

ChatGPT (OpenAI), a large language model (LLM) powered by artificial intelligence released in late 2022, may be a powerful tool for improving the availability of accurate and readable information for rare and complex diseases like amyloidosis. Unlike traditional search engines, ChatGPT generates human-like text in a conversational format through an intuitive user interface. This is achieved with reinforcement learning from human feedback, wherein the model's responses are refined through feedback loops to optimize responses [7]. With ongoing improvement and training using an extensive data set spanning diverse topics including medicine, ChatGPT's accuracy and reliability in answering questions are expected to improve.

Prior Work

Prior studies have demonstrated ChatGPT's impressive accuracy and reliability in answering clinical questions across multiple medical specialties [8-10]. One study found the model's generated responses were significantly higher in both quality and empathy compared to physicians when answering medical questions posted to social media, further bolstering the dynamic nature of this technology [11]. In March 2023, ChatGPT-4, the successor to ChatGPT-3.5, was released and has demonstrated superior performance in answering clinical questions across multiple fields of medicine [12-15]. In addition to accuracy and reliability, the readability of ChatGPT's responses is an active area of investigation. Several studies related to ophthalmology and endocrinology have revealed that responses by ChatGPT-4 often exceed the fifth- to sixth-grade reading level recommended by the American Medical Association (AMA) [16-18]. While the literature examining LLM responses to clinical questions is growing, studies examining rare diseases are limited. Furthermore, there are currently no studies examining ChatGPT's ability in answering questions related to amyloidosis.

Aims of This Study

As with any emerging technology, rigorous evaluation of these models' capabilities and limitations is essential to ensuring effective and safe implementation during their nascent stages before broad adoption by patients and providers. This study aims to build upon previous literature by using a multidisciplinary approach in assessing ChatGPT's (1) accuracy in answering questions related to amyloidosis, particularly concerning cardiology, gastroenterology, and neurology; (2) reproducibility of responses; (3) readability; and (4) comparison of performance between ChatGPT-4 and ChatGPT-3.5.

Methods

Question Curation

A total of 98 amyloidosis-related questions were sourced from the frequently asked questions section of websites for professional medical societies and institutions. Questions from amyloidosis Facebook support groups were also incorporated to represent a more comprehensive patient perspective. Of these questions, 56 addressed general amyloidosis topics, while 42 were specific to cardiology (n=12), gastroenterology (n=15), and neurology (n=15). Each question was inputted twice into ChatGPT-4 (version updated on March 14, 2023) and ChatGPT-3.5 (version updated on February 9, 2023) except for neurology-related questions, which were only inputted into ChatGPT-4 due to reviewer availability. At the time of data collection, ChatGPT-4 required a paid monthly subscription. Furthermore, the models were without internet access, and their training data were limited to information prior to September 2021.

Accuracy and Reproducibility

The accuracy of responses was assessed using the scale: (1) comprehensive, (2) correct but inadequate, (3) some correct and some incorrect, and (4) completely incorrect. Reproducibility was evaluated by categorizing each of the 2 responses of each question into those containing either no incorrect information (comprehensive and correct but inadequate) or those with incorrect information (some correct and some incorrect and completely incorrect). Questions that produced responses in different grading categories were deemed nonreproducible. Two independent reviewers (RG and DCK), board-certified in cardiology and gastroenterology with expertise in amyloidosis, assessed general amyloidosis questions and those of their respective specialties. Discrepancies in general question grading were resolved through discussion to reach a consensus. An additional reviewer (AAH), board-certified in neurology and specializing in amyloidosis, graded the neurology-specific responses for ChatGPT-4.

Readability

The readability of ChatGPT-4's responses was also assessed using the *Textstat* library in Python (Python Software Foundation) and the *Textstat readability* package in R software (R Foundation for Statistical Computing). The readability level was quantified either as a readability index or by using a predicted grade level, the latter indicating the US educational grade, at which the responses are comprehensible.

Statistical Analysis

Categorical variables were presented as counts and percentages, while continuous variables were presented as means and SDs. Bivariate analysis consisted of Fisher exact test for categorical variables. Microsoft Excel (version 16.68; Microsoft Corp) was used for all statistical analysis.

Ethical Considerations

Since all responses and outputs from ChatGPT were publicly available, approval from the institutional review board was not sought, and no informed consent was required.

Results

In this study, ChatGPT's responses were predominantly correct and also comprehensive (Table 1). Specifically, ChatGPT-4 (n=98) provided correct answers in 93 (95%) instances, with a notable 82 (84%) being graded as comprehensive. ChatGPT-3.5 (n=83) also performed well, delivering correct answers for 74 (89%) cases and comprehensive responses in 66 (79%) cases.

Table 1. Accuracy of responses by ChatGPT-3.5 and ChatGPT-4 to amyloidosis-related questions stratified by question subgroup.

Question subgroup	Responses, n (%)	
	ChatGPT-3.5	ChatGPT-4
Overall (n=83 for Chat GPT-3.5 and n=98 for Chat GPT-4)		
Comprehensive	66 (79)	82 (84)
Correct but inadequate	8 (10)	11 (11)
Some correct and some incorrect	8 (10)	5 (5)
Completely incorrect	1 (1)	0 (0)
General questions (n=56)		
Comprehensive	48 (86)	53 (95)
Correct but inadequate	4 (7)	3 (5)
Some correct and some incorrect	4 (7)	0 (0)
Completely incorrect	0 (0)	0 (0)
Cardiology questions (n=12)		
Comprehensive	10 (83)	10 (83)
Correct but inadequate	0 (0)	2 (17)
Some correct and some incorrect	2 (17)	0 (0)
Completely incorrect	0 (0)	0 (0)
Gastroenterology questions (n=15)		
Comprehensive	8 (53)	9 (60)
Correct but inadequate	4 (27)	3 (20)
Some correct and some incorrect	2 (13)	3 (20)
Completely incorrect	1 (7)	0 (0)
Neurology questions (n=15)		
Comprehensive	—[a]	10 (67)
Correct but inadequate	—	3 (20)
Some correct and some incorrect	—	2 (13)
Completely incorrect	—	0 (0)

[a]Not available.

When stratified by question category, both ChatGPT-4 and ChatGPT-3.5 excelled in general topics (n=56), where 53 (95%) and 48 (86%) of their responses, respectively, were comprehensive, though this difference was not statistically significant (*P*=.12). For cardiology, ChatGPT-4 was particularly accurate, correctly answering all 12 questions compared to ChatGPT-3.5's 10 (83%) responses (*P*=.48). In gastroenterology (n=15), both models produced correct responses for 80% (n=12) of questions. However, their comprehensiveness varied slightly with ChatGPT-3.5 at 8 (53%) and ChatGPT-4 at 9 (60%). In neurology (n=15), ChatGPT-4's responses were graded as comprehensive for 10 (67%).

Overall, ChatGPT-3.5 and ChatGPT-4 generated incorrect information in 9 of 83 (11%) and 5 of 98 (5%) responses, respectively. Notably, ChatGPT-3.5 produced 1 "completely incorrect" response regarding amyloidosis treatment of the gastrointestinal tract, involving the recommendation of probiotics and digestive enzymes (Multimedia Appendix 1). An example of a "some correct and some incorrect" response from ChatGPT-3.5 related to the management of atrial fibrillation in patients with amyloidosis. The model correctly described similar rate control and anticoagulation strategies for patients with amyloidosis having atrial fibrillation compared to those without amyloidosis but understated the prevalence of atrial fibrillation in ATTR. ChatGPT-4, on the other hand, did not produce any completely incorrect responses but did provide a response categorized as "correct but inadequate" by omitting autonomic symptoms in amyloidosis-related neuropathy. Regarding reproducibility, ChatGPT-4 showed a higher rate of 96 of 98 (98%) reproducible responses compared to 73 of 83 (88%) for ChatGPT-3.5 (Table 2).

Table 2. Reproducibility of responses by ChatGPT-3.5 and ChatGPT-4 to amyloidosis-related questions categorized by question subgroup.

Question subgroup	Responses, n (%)	
	ChatGPT-3.5	ChatGPT-4
Overall (n=83 for ChatGPT-3.5 and n=98 for ChatGPT-4)	73 (88)	96 (98)
General (n=56)	49 (88)	55 (98)
Cardiology (n=12)	10 (83)	12 (100)
Gastroenterology (n=15)	14 (93)	15 (100)
Neurology (n=15)	—[a]	14 (93)

[a]Not available.

In terms of readability, ChatGPT-4's responses varied but were consistently well above the AMA's recommended fifth- to sixth-grade reading level. The Flesch-Kincaid Grade Level scale rated them between a high school sophomore and a graduate level, averaging at a college level (mean 15.5, SD 1.9; range 10.3-21.7; Table 3). The Flesch Reading Ease scores, on a scale of 0 to 100, averaged at 23.3 (SD 9.4), indicating a college graduate level of complexity. Additional readability metrics showed a broad range of scores, all with similar advanced reading levels: Simple Measure of Gobbledygook (range 12.8-20.2), Gunning Fog Index (range 14.3-24.2), Coleman-Liau Index (range 10.5-18.3), Automated Readability Index (range 9.9-24.3), FORCAST Grade Level (range 10.3-13.4), and Powers Sumner Kearl Grade (range 6.8-9.4).

Table 3. Readability of responses by ChatGPT-4 to amyloidosis-related questions.

Readability metric	Score, mean (SD)	Range
Flesch Reading Ease	23.3 (9.4)	5.6-47.9
Flesch-Kincaid Grade Level	15.5 (1.9)	10.3-21.7
Simple Measure of Gobbledygook	16.7 (1.6)	12.8-20.2
Gunning Fog Index	19.1 (2.3)	14.3-24.2
Coleman-Liau Index	15.3 (1.4)	10.5-18.3
Automated Readability Index	15.6 (2.1)	9.9-24.3
FORCAST Grade Level	12.1 (0.53)	10.3-13.4
Powers Sumner Kearl Grade	8.2 (0.55)	6.8-9.4

Discussion

Principal Results

Literature examining ChatGPT's knowledge regarding rare diseases, such as amyloidosis, is limited compared to that of more prevalent health conditions. In this study, we employed an interdisciplinary panel of amyloidosis experts from cardiology, gastroenterology, and neurology to evaluate the accuracy and reproducibility of ChatGPT-4's and ChatGPT-3.5's responses to amyloidosis-related questions. Furthermore, the readability of responses by ChatGPT-4 was examined. ChatGPT-4 and ChatGPT-3.5 produced comprehensive responses to 53 (95%) and 48 (86%) general questions, respectively. Incorrect information was found in 5 of 98 (5%) and 9 of 83 (11%) responses from ChatGPT-4 and ChatGPT-3.5, respectively (P=.17), with 1 of 83 (1%) ChatGPT-3.5 responses graded as completely incorrect. The models also provided high reproducibility in accuracy of responses overall, with ChatGPT-4 and ChatGPT-3.5 generating 96 of 98 (98%) and 73 of 83 (88%) reproducible responses, respectively. However, the readability of ChatGPT-4's responses exceeded the AMA's recommended fifth- to sixth-grade reading level for PEMs, with readability at a college reading level on average.

Comparison With Prior Work

Previous studies have shown ChatGPT's impressive knowledge when assessing both common and rare diseases. The model has displayed extensive knowledge regarding cardiovascular disease prevention [8]. In more intricate scenarios such as clinical vignettes describing atrial fibrillation, congenital heart disease, and heart failure, its answers were assessed as predominantly reliable, valuable for patients, and crucially, not hazardous. Interestingly, many of these responses were favored over those generated by a standard Google search [19]. Similar results have been shown in several studies involving gastrointestinal-related topics such as cirrhosis, hepatocellular carcinoma, and bariatric surgery [9,10], with ChatGPT-4 demonstrating a significant improvement in knowledge compared to ChatGPT-3.5 [12,15]. Mehnen et al [13] demonstrated superior diagnostic precision of rare diseases by ChatGPT-4 compared to ChatGPT-3.5 as well. Our results showed comparable overall accuracy and reproducibility to previous studies, with both models generating consistent and reliable information. Although not meeting the level of significance as seen in prior research, ChatGPT-4 did generate fewer responses with incorrect information than ChatGPT-3.5 in this study.

The superior performance of ChatGPT-4 in prior studies may stem from multiple factors inherent to the design of each model. ChatGPT-4 was trained on a larger body of information, potentially exposing the model to a wider range of medical information. ChatGPT-4 has been reported to possess more advanced reasoning capabilities, allowing the model to better formulate explanations tailored to the input provided. Finally, the training of ChatGPT-4 may have provided the model with an advantage [14].

Limitations of ChatGPT

ChatGPT holds the potential to enhance clinical practice in the context of amyloidosis, but notable limitations exist. Chief among these is the undisclosed origin of ChatGPT's primary training data set, paired with its inability to regularly provide citations for its responses. Directly referencing established medical sources would bolster its clinical credibility. Moreover, ChatGPT sometimes produces responses referred to as "hallucinations," which are confident sounding, yet completely incorrect answers. The data set's scope is further limited to information prior to September 2021 [7]. The quality of responses generated by ChatGPT is affected by the nature of the prompts inputted by the user. Prompt engineering has been shown to significantly alter the models' output both in quality and comprehensiveness. Future studies would benefit from including the testing of different prompts and their effect on response output in the context of amyloidosis. Furthermore, concerted efforts in increasing patient and provider knowledge regarding prompt engineering may better facilitate the future effective use of these models. This study highlights the need for improvements in response readability to ensure equitable use of this technology across all patient populations. Similarly, other studies involving hypothyroidism in the setting of pregnancy and retinal surgery have also noted ChatGPT to produce information at a college reading level and beyond [17,18]. Furthermore, the majority of studies in the literature have examined the model's performance in English, with a limited body of literature examining non-English languages [20-22]. More studies are needed to ensure the optimization of model performance across a wide range of languages.

Ethical Implications

Beyond model-specific challenges, ethical issues remain unresolved. Potential biases introduced during training could skew user outputs. Clinical research bias, such as the overrepresentation of White populations [23], might also persist within the model. There is a growing body of literature examining implicit bias in responses from LLMs with conflicting results [24-26]. Equitable access is another concern; lower socioeconomic groups might face barriers in accessing such technology due to hardware and internet constraints. Privacy is a further point of contention, though OpenAI's option to disable chat history storage addresses some concerns [27]. Regulatory oversight, as suggested by the Food and Drug Administration, is paramount. The proposed regulation would align artificial intelligence health care tools with medical device standards, emphasizing repeated validation and testing at each stage of development [28]. Additionally, physician panels should advise technical developers, ensuring patient safety and prioritizing equitable, outcome-driven patient care.

Strengths and Limitations of This Study

This study's strengths include being among the first in using a multidisciplinary approach to evaluate ChatGPT's knowledge of amyloidosis. This holistic approach enabled a thorough assessment of ChatGPT's abilities in addressing clinical queries related to amyloidosis, a rare disease necessitating advancements in health education, diagnostics, and management for improved patient outcomes. However, this study is not without its limitations. We relied on a single physician reviewer for specialty-specific responses, which is subjective and prone to bias. Research could bolster validity by engaging multiple reviewers within each specialty to minimize the potential for subjective bias. It would also be beneficial to include physicians specializing in hematology, oncology, and nephrology as reviewers due to their integral involvement in caring for patients with amyloidosis. Furthermore, we recommend including patients and all members of the health care team when reviewing the quality of responses. While we took a systematic approach when curating questions, our list may not comprehensively represent all potential patient questions related to amyloidosis.

Conclusions

ChatGPT delivered accurate and reliable responses to amyloidosis-related questions across general and specialty-specific questions. ChatGPT has the potential to serve as a supplemental tool in disseminating vital health education to patients in the future. However, the presence of some incorrect responses underscores the necessity of continued improvements and fine-tuning of future iterations prior to incorporation into clinical practice. Furthermore, improvement in the readability of responses is essential to ensuring equal access to this technology by all patients. We advocate for the use of this technology as an adjunct and not a replacement to care and advice provided by licensed health care professionals. In its current state, there are also limitations and ethical concerns that need to be resolved before the technology may be widely implemented in health care in a safe and equitable manner.

Acknowledgments

ChatGPT-4, the version updated on March 14, 2023, by OpenAI was used in the final editing process of this paper to improve readability.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

RG is a consultant for Pfizer, Alnylam, and AstraZeneca. None of the other authors have interests to disclose.

Multimedia Appendix 1
Examples of prompts with corresponding ChatGPT responses and reviewer accuracy grades.
[DOCX File , 19 KB - cardio_v8i1e53421_app1.docx]

References

1. Papingiotis G, Basmpana L, Farmakis D. Cardiac amyloidosis: epidemiology, diagnosis and therapy. E-Journal Cardiol Pract 2021;19:19-21 [FREE Full text]
2. Bajwa F, O'Connor R, Ananthasubramaniam K. Epidemiology and clinical manifestations of cardiac amyloidosis. Heart Fail Rev 2022;27(5):1471-1484. [doi: 10.1007/s10741-021-10162-1] [Medline: 34694575]
3. Kharoubi M, Bézard M, Galat A, Le Bras F, Poullot E, Molinier-Frenkel V, et al. History of extracardiac/cardiac events in cardiac amyloidosis: prevalence and time from initial onset to diagnosis. ESC Heart Fail 2021;8(6):5501-5512 [FREE Full text] [doi: 10.1002/ehf2.13652] [Medline: 34714605]
4. Kittleson MM, Ruberg FL, Ambardekar AV, Brannagan TH, Cheng RK, Clarke JO, et al. 2023 ACC expert consensus decision pathway on comprehensive multidisciplinary care for the patient with cardiac amyloidosis: a report of the American College of Cardiology Solution Set Oversight Committee. J Am Coll Cardiol 2023;81(11):1076-1126 [FREE Full text] [doi: 10.1016/j.jacc.2022.11.022] [Medline: 36697326]
5. Falcão M, Allocca M, Rodrigues AS, Granjo P, Francisco R, Pascoal C, et al. A community-based participatory framework to co-develop Patient Education Materials (PEMs) for rare diseases: a model transferable across diseases. Int J Environ Res Public Health 2023;20(2):968 [FREE Full text] [doi: 10.3390/ijerph20020968] [Medline: 36673723]
6. Talking points about health literacy. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/healthliteracy/shareinteract/TellOthers.html# [accessed 2024-03-01]
7. ChatGPT: optimizing language models for dialogue. OpenAI. 2023. URL: https://online-chatgpt.com [accessed 2023-02-18]
8. Sarraju A, Bruemmer D, Van Iterson E, Cho L, Rodriguez F, Laffin L. Appropriateness of cardiovascular disease prevention recommendations obtained from a popular online chat-based artificial intelligence model. JAMA 2023;329(10):842-844 [FREE Full text] [doi: 10.1001/jama.2023.1044] [Medline: 36735264]
9. Yeo YH, Samaan JS, Ng WH, Ting PS, Trivedi H, Vipani A, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. Clin Mol Hepatol 2023;29(3):721-732 [FREE Full text] [doi: 10.3350/cmh.2023.0089] [Medline: 36946005]
10. Samaan JS, Yeo YH, Rajeev N, Hawley L, Abel S, Ng WH, et al. Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery. Obes Surg 2023;33(6):1790-1796 [FREE Full text] [doi: 10.1007/s11695-023-06603-5] [Medline: 37106269]
11. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
12. Yeo YH, Samaan JS, Ng WH, Ma X, Ting PS, Kwak MS, et al. GPT-4 outperforms ChatGPT in answering non-English questions related to cirrhosis. medRxiv Preprint posted online on May 5, 2023 [FREE Full text] [doi: 10.1101/2023.05.04.23289482]
13. Mehnen L, Gruarin S, Vasileva M, Knapp B. ChatGPT as a medical doctor? A diagnostic accuracy study on common and rare diseases. medRxiv Preprint posted online on April 27, 2023 [FREE Full text] [doi: 10.1101/2023.04.20.23288859]
14. OpenAI. GPT-4 technical report. ArXiv Preprint posted online on December 19, 2023https://cdn.openai.com/papers/gpt-4.pdf [FREE Full text]
15. King RC, Bharani V, Shah K, Yeo YH, Samaan JS. GPT-4V passes the BLS and ACLS examinations: an analysis of GPT-4V's image recognition capabilities. Resuscitation 2024;195:110106 [FREE Full text] [doi: 10.1016/j.resuscitation.2023.110106] [Medline: 38160904]
16. Weis BD. Health literacy: a manual for clinicians. American Medical Association Foundation and American Medical Association. 2003. URL: http://lib.ncfh.org/pdfs/6617.pdf [accessed 2024-03-01]
17. Onder CE, Koc G, Gokbulut P, Taskaldiran I, Kuskonmaz SM. Evaluation of the reliability and readability of ChatGPT-4 responses regarding hypothyroidism during pregnancy. Sci Rep 2024;14(1):243 [FREE Full text] [doi: 10.1038/s41598-023-50884-w] [Medline: 38167988]
18. Momenaei B, Wakabayashi T, Shahlaee A, Durrani AF, Pandit SA, Wang K, et al. Appropriateness and readability of ChatGPT-4-generated responses for surgical treatment of retinal diseases. Ophthalmol Retina 2023;7(10):862-868. [doi: 10.1016/j.oret.2023.05.022] [Medline: 37277096]
19. Van Bulck L, Moons P. What if your patient switches from Dr. Google to Dr. ChatGPT? A vignette-based survey of the trustworthiness, value, and danger of ChatGPT-generated responses to health questions. Eur J Cardiovasc Nurs 2024;23(1):95-98 [FREE Full text] [doi: 10.1093/eurjcn/zvad038] [Medline: 37094282]

20. Samaan JS, Yeo YH, Ng WH, Ting PS, Trivedi H, Vipani A, et al. ChatGPT's ability to comprehend and answer cirrhosis related questions in Arabic. Arab J Gastroenterol 2023;24(3):145-148. [doi: 10.1016/j.ajg.2023.08.001] [Medline: 37673708]
21. Guillen-Grima F, Guillen-Aguinaga S, Guillen-Aguinaga L, Alas-Brun R, Onambele L, Ortega W, et al. Evaluating the efficacy of ChatGPT in navigating the Spanish Medical Residency Entrance Examination (MIR): promising horizons for AI in clinical medicine. Clin Pract 2023;13(6):1460-1487 [FREE Full text] [doi: 10.3390/clinpract13060130] [Medline: 37987431]
22. Panthier C, Gatinel D. Success of ChatGPT, an AI language model, in taking the French language version of the European Board of Ophthalmology examination: a novel approach to medical knowledge assessment. J Fr Ophtalmol 2023;46(7):706-711 [FREE Full text] [doi: 10.1016/j.jfo.2023.05.006] [Medline: 37537126]
23. Lolic M, Araojo R, Okeke M, Woodcock J. Racial and ethnic representation in US clinical trials of new drugs and biologics, 2015-2019. JAMA 2021;326(21):2201-2203 [FREE Full text] [doi: 10.1001/jama.2021.16680] [Medline: 34874429]
24. Zack T, Lehman E, Suzgun M, Rodriguez JA, Celi LA, Gichoya J, et al. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. Lancet Digit Health 2024;6(1):e12-e22 [FREE Full text] [doi: 10.1016/S2589-7500(23)00225-X] [Medline: 38123252]
25. Hanna JJ, Wakene AD, Lehmann CU, Medford RJ. Assessing racial and ethnic bias in text generation for healthcare-related tasks by ChatGPT. medRxiv Preprint posted online on August 28, 2023 [FREE Full text] [doi: 10.1101/2023.08.28.23294730] [Medline: 37693388]
26. Omiye A, Lester JC, Spichak S, Rotemberg V, Daneshjou R. Large language models propagate race-based medicine. NPJ Digit Med 2023;6(1):195 [FREE Full text] [doi: 10.1038/s41746-023-00939-z] [Medline: 37864012]
27. New ways to manage your data in ChatGPT. OpenAI. URL: https://openai.com/blog/new-ways-to-manage-your-data-in-chatgpt [accessed 2023-08-15]
28. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD). US Food and Drug Administration. 2019. URL: https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf [accessed 2024-02-28]

Abbreviations

AMA: American Medical Association
ATTR: transthyretin amyloidosis
LLM: large language model
PEM: patient education material

Edited by A Mavragani; submitted 10.10.23; peer-reviewed by D Chrimes, L Zhu; comments to author 10.12.23; revised version received 31.01.24; accepted 22.02.24; published 19.04.24.

Please cite as:
King RC, Samaan JS, Yeo YH, Peng Y, Kunkel DC, Habib AA, Ghashghaei R
A Multidisciplinary Assessment of ChatGPT's Knowledge of Amyloidosis: Observational Study
JMIR Cardio 2024;8:e53421
URL: https://cardio.jmir.org/2024/1/e53421
doi:10.2196/53421
PMID:38640472

Original Paper

Use of Machine Learning for Early Detection of Maternal Cardiovascular Conditions: Retrospective Study Using Electronic Health Record Data

Nawar Shara[1,2], PhD; Roxanne Mirabal-Beltran[3], PhD; Bethany Talmadge[4], BS; Noor Falah[1], MS; Maryam Ahmad[1], BS; Ramon Dempers[4], MS; Samantha Crovatt[4], BS; Steven Eisenberg[4], MD; Kelley Anderson[3], PhD

[1]MedStar Health Research Institute, Hyattesville, MD, United States

[2]Georgetown-Howard Universities Center for Clinical and Translational Science, Washington, DC, DC, United States

[3]School of Nursing, Georgetown University, Washington, DC, United States

[4]Invaryant Inc, Roswell, GA, United States

Corresponding Author:
Roxanne Mirabal-Beltran, PhD
School of Nursing
Georgetown University
3700 Reservoir Road, NW
Washington, DC, 20057
United States
Phone: 1 2026873496
Email: rm1910@georgetown.edu

Abstract

Background: Cardiovascular conditions (eg, cardiac and coronary conditions, hypertensive disorders of pregnancy, and cardiomyopathies) were the leading cause of maternal mortality between 2017 and 2019. The United States has the highest maternal mortality rate of any high-income nation, disproportionately impacting those who identify as non-Hispanic Black or Hispanic. Novel clinical approaches to the detection and diagnosis of cardiovascular conditions are therefore imperative. Emerging research is demonstrating that machine learning (ML) is a promising tool for detecting patients at increased risk for hypertensive disorders during pregnancy. However, additional studies are required to determine how integrating ML and big data, such as electronic health records (EHRs), can improve the identification of obstetric patients at higher risk of cardiovascular conditions.

Objective: This study aimed to evaluate the capability and timing of a proprietary ML algorithm, Healthy Outcomes for all Pregnancy Experiences-Cardiovascular-Risk Assessment Technology (HOPE-CAT), to detect maternal-related cardiovascular conditions and outcomes.

Methods: Retrospective data from the EHRs of a large health care system were investigated by HOPE-CAT in a virtual server environment. Deidentification of EHR data and standardization enabled HOPE-CAT to analyze data without pre-existing biases. The ML algorithm assessed risk factors selected by clinical experts in cardio-obstetrics, and the algorithm was iteratively trained using relevant literature and current standards of risk identification. After refinement of the algorithm's learned risk factors, risk profiles were generated for every patient including a designation of standard versus high risk. The profiles were individually paired with clinical outcomes pertaining to cardiovascular pregnancy conditions and complications, wherein a delta was calculated between the date of the risk profile and the actual diagnosis or intervention in the EHR.

Results: In total, 604 pregnancies resulting in birth had records or diagnoses that could be compared against the risk profile; the majority of patients identified as Black (n=482, 79.8%) and aged between 21 and 34 years (n=509, 84.4%). Preeclampsia (n=547, 90.6%) was the most common condition, followed by thromboembolism (n=16, 2.7%) and acute kidney disease or failure (n=13, 2.2%). The average delta was 56.8 (SD 69.7) days between the identification of risk factors by HOPE-CAT and the first date of diagnosis or intervention of a related condition reported in the EHR. HOPE-CAT showed the strongest performance in early risk detection of myocardial infarction at a delta of 65.7 (SD 81.4) days.

Conclusions: This study provides additional evidence to support ML in obstetrical patients to enhance the early detection of cardiovascular conditions during pregnancy. ML can synthesize multiday patient presentations to enhance provider decision-making and potentially reduce maternal health disparities.

(JMIR Cardio 2024;8:e53091) doi:10.2196/53091

KEYWORDS

machine learning; preeclampsia; cardiovascular; maternal; obstetrics; health disparities; woman; women; pregnancy; pregnant; cardiovascular; cardiovascular condition; retrospective study; electronic health record; EHR; technology; decision-making; health disparity; virtual server; thromboembolism; kidney failure; HOPE-CAT

Introduction

All other high-income nations in the world have substantially lower maternal mortality rates compared to the United States [1]. Maternal mortality rates in the United States have increased over the past 30 years [1,2], increasing approximately 85% between 2018 and 2021 [3]. There were 23.8 and 32.9 maternal deaths in the US per 100,000 live births in 2020 and 2021, respectively [4], and although the United States has one of the highest health care spending rates [5], the projected trends of maternal mortality are anticipated to continue to rise. Racial and ethnic disparities in maternal mortality rates have not only persisted, but differences in rates have widened. Non-Hispanic Black (Black) women are significantly more likely to die of pregnancy-related causes than non-Hispanic White (White) women (69.9 and 26.6 deaths per 100,000 live births, respectively) [4]. Cardiovascular conditions (eg, cardiac and coronary conditions, hypertensive disorders of pregnancy, and cardiomyopathies) were the leading cause of maternal mortality between 2017 and 2019 (27.8%) [6]. Black women have higher rates of cardiovascular morbidity and mortality than women of other races and ethnicities [7].

Programs across the United States are integrating technology to mitigate increasing maternal mortality rates, and there is a growing body of literature reporting the use of machine learning (ML) to identify patients at increased risk for hypertensive disorders in pregnancy [8,9]. Hoffman et al [8] used ML to accurately predict maternal readmission due to the complications of hypertensive disorders. ML models have also shown promising results for predicting maternal risk of hypertensive disorders of pregnancy and other cardiovascular conditions [10-12], including improved prediction accuracy when gestational age, epidemiology, hemodynamics, and biochemistry data are incorporated [9]. However, a fundamental gap remains to identify pregnant individuals at higher risk of morbidity and mortality using the power of ML and big data including electronic health records (EHRs). Therefore, this study evaluated ML technology, specifically Invaryant's Healthy Outcomes for all Pregnancy Experiences-Cardiovascular-Risk Assessment Technology (HOPE-CAT), applied to a large health care system database, to enable early and effective risk identification of cardiovascular conditions associated with complications in pregnancy, including maternal morbidity and mortality.

Methods

Overview

This study was conducted using retrospective data of 32,409 obstetric patients sourced from the EHR of a large, US-based health care system with a documented birth between January 1, 2017, and December 31, 2020. This timeline was chosen to capture as many pregnancies as possible while considering factors related to when various EHR systems within MedStar Health were implemented.

Sample

For this study, patients were selected if they met inclusion criteria: patients between the ages of 18 and 40 years at the time of the index pregnancy–related visit and who had more than 1 pregnancy-related medical encounter. To ensure the richness of the data and adequate data points to test the ability of HOPE-CAT to assess a diagnosed condition in the EHR, patients were excluded if they had limited EHR data or a single pregnancy–related medical encounter (n=11,485) and if the initial encounter within the EHR was a birth with no corresponding pregnancy data (n=14,855). This resulted in a sample of 6069 patients for analysis.

HOPE-CAT

The ML risk assessment algorithm, HOPE-CAT, was trained via causal inference to analyze patient data and identify factors associated with the development of conditions leading to complications during pregnancy and postbirth. Clinical criteria were designated by clinical experts based on relevant literature and current standards in evaluating risk during pregnancy. Following an iterative training process, using anonymized EHRs, patient records were reviewed by HOPE-CAT on a simulated encounter-by-encounter basis. HOPE-CAT then surfaced and refined the most frequent indicators of risk within this patient population. Data analysis then validated findings against relevant data, and clinical experts then reviewed the training results and made necessary adjustments before final manual testing was conducted, and HOPE-CAT was deployed for this study. The risk factors assessed by HOPE-CAT are divided into 2 categories—static and variable (Textbox 1). Static risk factors are defined as those characteristics that do not change or change infrequently such as race, ethnicity, age, medical history, and family history. Variable risk factors are defined as characteristics that change more frequently, such as blood pressure, heart rate, and symptoms such as headache and shortness of breath.

Risk profiles were generated by patient encounters and included basic patient demographics (eg, patient ID number, age at visit, and race) and any surfaced risk factors (Textbox 1) that were noted in the patient's record at the time of appointment. HOPE-CAT generated 2 types of risk profiles that quantify the risk factors recorded or identified in each visit: standard and high risk. High-risk, or *red flag*, profiles are defined as specific severe indicators or having 4 or more signs of risk present in a single encounter. Severe indicators that trigger a *red flag*, indicating greater risk compared to other parameters, include resting heart rate ≥120 bpm, systolic blood pressure ≥160 mm Hg, respiratory rate ≥30, oxygen saturation ≤94%, dyspnea, and orthopnea.

With this study, we primarily focused on validating the ability of the ML algorithm (HOPE-CAT) to identify pregnant individuals at high risk of cardiovascular disease within a retrospective data set. Therefore, the study design prioritized tracing outcomes for individuals flagged by the algorithm to assess its accuracy in predicting risk compared to diagnoses documented in medical records. For more extensive technical detail related to the development of HOPE-CAT, including data extraction, specifications of the computational systems, the development of the data set, and access to relevant systems, please see our previous work, Early Identification of Maternal Cardiovascular Risk Through Sourcing and Preparing Electronic Health Record Data: Machine Learning Study [10].

Textbox 1. Risk categories for maternal cardiovascular conditions; categories were divided into variable and static risk.

Symptoms (variable risk): risk factors with no measurement scale specified were identified from the data by diagnosis codes as yes (present) or no (absent)

- Asthma, unresponsive to therapy
- Chest pain
- Dizziness or syncope
- Dyspnea (red flag risk)
- Headache, new or worsening
- Heart palpitations
- Orthopnea (red flag risk)
- Swelling of face or hands
- Tachypnea

Physical findings (variable risk): risk factors with no measurement scale specified were identified from the data by diagnosis codes as yes (present) or no (absent)

- Basilar crackles in lungs
- Loud heart murmur
- Oxygen saturation ≤96% (≤94% considered a red flag risk)
- Respiratory rate ≥ 24 (≥30 considered a red flag risk)
- Resting heart rate ≥ 110 beats per minute (≥ 120 beats per minute considered a red flag risk)
- Systolic blood pressure ≥ 140 mm Hg (≥160 mm Hg considered a red flag risk)

Medical history (static risk)

- Age (continuous in years)
- Chronic hypertension existing prior to pregnancy
- Ethnicity
- History of chemotherapy
- History of complications in labor or birth
- History of heart disease
- Prepregnancy obesity (BMI≥35)
- Pregestational diabetes
- Race
- Substance use (eg, nicotine, cocaine, alcohol, and methamphetamines)

Ethical Considerations

Ethics approval was obtained from the Georgetown-MedStar institutional review board (STUDY00003534) and adhered to all appropriate ethical reviews and approvals, as per institutional guidelines. Institutional review board approval covered secondary analysis without additional consent. Data access, extraction, transfer, and anonymization procedures were reviewed by the institution's data security team to ensure all necessary security requirements were implemented before the release of data. Data were deidentified before being transferred for analysis. All personal health information were removed, and patients were assigned a unique code to prevent reidentification; study identification numbers were known by a single member of the data team at the EHR institution to allow for any revalidation after original extraction. Events were also

deidentified to remove any comments or free-text entry fields that could potentially be identifiable. Adherence to institutional privacy and security policies ensured patient data were protected and secured throughout the project. As this was a secondary analysis, no compensation was provided.

Analysis

Within a virtual server environment created for this study, we systematically cleaned and standardized the EHR data (eg, organized and matched fields across tables and databases) before the deployment of an analysis by HOPE-CAT [10]. While the data set was sourced within a single health care system, standardization was necessary as data were sourced from multiple EHRs. Variables, however, were similarly defined across patients (eg, race, ethnicity, and preeclampsia). The EHR data were deidentified before analyses, enabling HOPE-CAT to truly analyze the data without consideration of race, ethnicity, or other variables that may lead to bias. Race, ethnicity, medical history, and family history were self-reported by patients and defined as recorded in the EHR.

HOPE-CAT was deployed to analyze patient data on an encounter-by-encounter basis, as in each visit and data collection on record were analyzed in the order of entry. Data from each encounter included patient demographics, physical findings, symptoms, and medical history and were analyzed by HOPE-CAT to identify changes and trends. In instances where HOPE-CAT detected any sign of risk, based on the training criteria, a risk profile was generated for the given patient and encounter [10]. Due to the nature of the available data, it was not possible to determine whether a patient's pregnancy was their first. Each pregnancy was analyzed in isolation, and risk profiles were only compared to outcomes within that individual pregnancy.

Following full analysis by HOPE-CAT, patient risk profiles were linked to any outcomes of cardiovascular pregnancy conditions and complications, specifically preeclampsia and eclampsia, cardiomyopathy, myocardial infarction (MI), heart failure, acute kidney disease and failure, cerebral infarction, pulmonary embolism, venous thromboembolism, and HELLP (hemolysis, elevated liver enzymes, and low platelets) syndrome. Once a risk profile was identified and linked with a patient's outcomes as stated in the EHR, a difference was calculated based on the date of the risk profile compared to the date of diagnosis or intervention recorded in the patient's records. For example, if HOPE-CAT detected a risk that indicated a patient may be experiencing symptoms of preeclampsia on day 114 and the patient's EHR data showed a diagnosis of preeclampsia on day 142, the difference would be 28 days. In-depth manual reviews were conducted by retrieving the relevant data against the results of HOPE-CAT and then cross-checking and tabulating each result. The tabulated information was further cross-checked by the independent quality team.

Results

Of the 6069 patients analyzed by HOPE-CAT, 5238 patients had 1 or more risk factors (Figure 1). A total of 1716 (32.8%) *red flag* risk profiles were identified among patients with 1 or more risk factors. Of the 1716 *red flag* risk profiles developed, HOPE-CAT identified risk profiles for 620 patients who could be matched with outcomes of interest in the patient's diagnosis and intervention (medication) data. There were patients in the final subset for which HOPE-CAT identified risk of cardiovascular conditions after a diagnosis was made (ie, the delta was a negative number). These patients were included for accurate representation.

Following the identification of all resulting risk profiles, 16 patients who were identified by HOPE-CAT in duplicate (eg, risks that tracked to 2 different outcomes) were combined. This resulted in a final sample of 604 patients for whom available records of diagnoses or interventions could be compared against and linked with the identified risk profiles.

The majority of the final sample self-identified as Black (n=482, 79.8%) and between the ages of 20 and 34 years (n=509, 84.4%; Table 1). Twenty-one patients had 2 pregnancies recorded during the study period.

Preeclampsia was the most common condition diagnosed in our sample (n=547, 90.6%), followed by thromboembolism (n=16, 2.7%) and acute kidney disease or failure (n=13, 2.2%; Table 2).

The average delta for the final subset of 604 patients was 56.8 (SD 69.7) days between when HOPE-CAT pinpointed risk factors during a patient's visit and the first date of diagnosis or intervention of a related condition in the patient's record. For patients who were diagnosed with preeclampsia, HOPE-CAT identified risk factors an average of 60.2 (SD 90.9) days earlier than the time point indicated in the records. Patients with a diagnosis of MI were identified with risk factors for MI on average 65.7 (SD 81.4) days earlier than the first reported date of diagnosis. For patients whose pregnancy experience resulted in cerebral infarction (n=4, 0.7%), the delta was 42.3 days. Of these 604 patients, 19 (3.1%) experienced 2 or more tracked conditions (eg, preeclampsia with acute kidney failure and peripartum cardiomyopathy).

Figure 1. Flow diagram showing sample selection process for HOPE-CAT (Healthy Outcomes for all Pregnancy Experiences-Cardiovascular-Risk Assessment Technology) analysis.

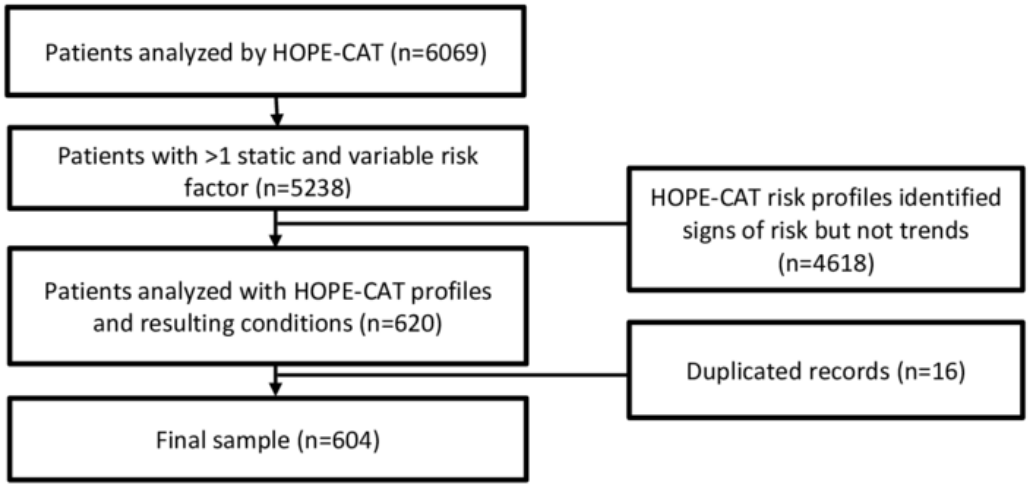


Table 1. Participant demographics of race, ethnicity, and age in years (N=604).

Characteristic	Values, n (%)
Race[a]	
American Indian or Alaska Native	1 (0.2)
Asian	1 (0.2)
Black	482 (79.8)
White	86 (14.2)
None of the above or unknown race	34 (5.6)
Ethnicity	
Hispanic	7 (1.2)
Non-Hispanic	597 (98.8)
Age (years)	
18-19	1 (0.17)
20-34	509 (84.4)
35-40	94 (15.6)

[a]Self-identification as reported in the electronic health record.

Table 2. Patient subset by condition identified in the electronic health record, including delta (N=604).

Conditions	Patients, n (%[a])	Delta mean (SD) (days)
Preeclampsia	547 (90.6)	60.2 (90.9)
Thromboembolism	16 (2.7)	15.3 (37.7)
Acute kidney disease and failure	13 (2.2)	25.1 (43.2)
Cardiomyopathy	10 (1.7)	13.9 (43.9)
Eclampsia	9 (1.3)	46.2 (58.3)
HELLP[b] syndrome	7 (1.2)	34.0 (34.8)
Heart failure	5 (0.8)	13.6 (13.2)
Cerebral infarction	4 (0.7)	42.3 (36.8)
Myocardial infarction	3 (0.5)	65.7 (81.4)

[a]Percentage of patients does not add to 100 as some patients had multiple conditions.

[b]HELLP: hemolysis, elevated liver enzymes, and low platelets.

Discussion

Principal Findings

In this study, high-risk cardiovascular conditions in pregnancy were identified earlier than determined by a health care provider by leveraging Invaryant's HOPE-CAT ML technology to surface signals and trends in patients' medical records. HOPE-CAT enabled early and effective screening of potential risks from cardiovascular disease an average of 56.8 (SD 69.7) days earlier than the first date of diagnosis or intervention of a related condition as documented in the EHR.

Cardiovascular disease is the leading cause of death for women in the United States, and overall, US women experience challenges in cardiovascular care, with notable underdiagnosis and late presentations, making it less likely for them to receive appropriate treatments [13]. Black women, specifically, experience cardiovascular disease at a higher rate than White women [14,15] and are also significantly more likely to die of pregnancy-related causes [4]. Pregnancy-related disorders are not only associated with complications during pregnancy, but they also portend future cardiometabolic and long-term cardiovascular-related morbidity [13]. Implicit racial biases contribute to these health inequities, resulting in increased maternal morbidity and mortality when Black women with valid and important health concerns are dismissed [16]. Enabling early and effective screening for pre-existing comorbidities and early identification of risk with enhanced technological applications, like HOPE-CAT, independent of patient characteristics and descriptors or provider bias [17], has the potential to mitigate factors leading to racial biases [18]. Screening is imperative to promote optimal cardiovascular health to generate appropriate referrals and collaborations among health care specialties [19]. Cardio-obstetrics teams have demonstrated some promising outcomes in women with known cardiovascular disease, and early identification may facilitate the inclusion of patients who are pregnant into these multidisciplinary care teams in a timely manner [20].

HOPE-CAT also offers an important clinical tool for providers to enhance the early detection and intervention of cardiovascular disease in pregnancy by facilitating the delivery of care based on trends and synthesis of data over time rather than only the current presentation of the patient at a single visit. The ability to create risk profiles dependent on patient presentation and independent of provider recall may improve the accuracy of disease identification and promote changes in provider recommendations, monitoring, treatment, referrals, and even patients' self-monitoring and awareness of risk. Importantly, HOPE-CAT generates a risk profile based on multiple factors that are routinely collected in prenatal visits. This approach has been argued to be superior to prediction models that require information not routinely collected [21] or to "static and single-class conventional prediction methods" in the detection of hypertensive disorders in pregnancy and post partum [9]. A real benefit of HOPE-CAT lies in its deployment in areas with limited resources and providers without obstetrical specializations. Early identification of pregnant individuals at high risk for cardiovascular disease would allow more prompt referral to high-risk clinics and specialist care.

Finally, identifying and treating cardiovascular diseases early in a pregnancy can alleviate stress on health care systems including decreasing costs of hospitalization and urgent care, and more importantly, decreasing the risk of morbidity and mortality among expecting mothers. Maternal mortality affects a country's economic well-being. The total cost of US maternal mortality in 2019 was estimated to be US $32.3 billion from birth to the child reaching their fifth birthday [22]. In the United States, maternal mortality has been rising since 2000, even with respect to gross domestic product and health expenditure per capita [23]. It should be noted that while ML technology is a valuable tool for providers in the care of birthing persons resulting in decreased costs in the long term, the possibility exists that there could be additional short-term costs related to testing and interventions [8].

Comparison With Prior Work

Our retrospective study adds to the body of research exploring the use of ML in clinical practice by leveraging the power of EHR data to evaluate HOPE-CAT's early identification of cardiovascular risk. This study also adds to the limited ML research in the field of obstetrics. A review conducted between 2000 and 2018 including 386 studies reporting on the use of ML in clinical practice found only 10 studies focused specifically on the field of obstetrics and gynecology [24]. Another systematic review analyzed publications on the use of ML application in obstetrics and gynecology core discipline journals and found only 19 publications between 2000 and 2020 [25].

In obstetrics, ML has successfully been used to determine the clinical parameters most useful for predicting preeclampsia and hypertensive disorders of pregnancy [12,26,27]. Our study provides additional evidence to support the use of ML in the field of obstetrics. Unlike prior work, our purpose was not to determine which parameters were most predictive but rather how soon HOPE-CAT could determine risk, based on an iterative, encounter-by-encounter basis [28]. Specifically, this study offers a novel tool for monitoring cardiovascular risk during pregnancy using real-time trends, thereby assisting health care professionals in the provision of perinatal care for high-risk patients. Consistent with earlier inquiries on the deployment of ML within obstetrics, our ML model has undergone technical validation [10]. Nonetheless, the imperative remains for further research to establish the clinical validity of the model.

Limitations

This study had several limitations. Complete EHRs were not always available for proper analysis and we did not have complete access to patients' full health history [29]. Compared other studies applying ML to big data [27], we relied on EHRs and did not include other types of unstructured data (eg, clinical notes) in this analysis. This may have created limitations when interpreting results. In order to optimize the performance of HOPE-CAT, several steps were required to clean and standardize the data set due to sourcing from multiple EHRs. This does not negate the validity or capability of the ML

technology but does describe the nonstandardized, noninteroperable nature of EHRs data in the present day. The retrospective nature of our study was limited to data available within the health system's EHRs and has been identified as a limitation in many prior studies using ML [24]. However, the advantage of retrospective data is that it allows for in-depth manual reviews. The data for this study were collected between January 1, 2017, and December 31, 2020. An overlap with the COVID-19 pandemic may have resulted in skewed data; pregnant individuals may have been reluctant to seek care or faced additional barriers to accessing care in 2020, which may have resulted in delayed diagnosis of complications. Additionally, analysis by race, ethnicity, or age was not possible with our sample, and thus we may have failed to capture important differences, such as whether the average delta may have been different by ethnicity, race, or age. Finally, this study was conducted within 1 health care system, and the results may not be generalizable to a different demographic or setting.

Conclusions

The findings from this study provide the foundation for future work to evaluate ML prospectively, in vivo, in the real-world setting, and longitudinally during current pregnancies and to inform future pregnancies, postpartum events, and overall cardiovascular health. Future ML may integrate multiple data sources including unstructured data using natural language processing, wearables, remote patient monitoring devices (eg, blood pressure), and symptom surveys. Additionally, the integration of factors related to social determinants of health may inform solutions to advance health equity and address the increasing US maternal mortality rates that show widening disparities associated with race and ethnicity. To facilitate the reversal of this trend, it is imperative that risk identification occurs earlier in the pregnancy trajectory to allow for increased monitoring and referral to more specialized care [30] and that ML technology is leveraged to support maternal health screening in routine appointments. The results from this study of ML through HOPE-CAT provide foundational evidence to develop solutions to mitigate the harmful impacts of pregnancy and improve maternal health for all.

Acknowledgments

We would like to acknowledge the following individuals who contributed to this work and manuscript: Sarah McGaughey and John Katial. RMB was supported in part by the National Center for Advancing Translational Sciences of the National Institutes of Health under award KL2TR001432. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

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

Conflicts of Interest

None declared.

References

1. Aspen Health Strategy Group. Reversing the U.S. maternal mortality crisis. The Aspen Institute. 2021. URL: https://www.aspeninstitute.org/wp-content/uploads/2021/04/Maternal-Morality-Report.pdf [accessed 2024-03-23]
2. MacDorman MF, Declercq E, Cabral H, Morton C. Recent increases in the U.S. maternal mortality rate: disentangling trends from measurement issues. Obstet Gynecol 2016;128(3):447-455. [doi: 10.1097/AOG.0000000000001556]
3. Maternal health: outcomes worsened and disparities persisted during the pandemic. Government Accountability Office-23-105871. URL: https://www.gao.gov/assets/gao-23-105871.pdf [accessed 2023-03-24]
4. Hoyert DL. Maternal mortality rates in the United States, 2021. NCHS Health E-Stats. 2023. URL: https://www.cdc.gov/nchs/data/hestat/maternal-mortality/2021/maternal-mortality-rates-2021.pdf [accessed 2024-03-23]
5. U.S. health care spending highest among developed countries. Johns Hopkins Bloomberg School of Public Health. URL: https://publichealth.jhu.edu/2019/us-health-care-spending-highest-among-developed-countries [accessed 2022-11-02]
6. Trost S, Beauregard J, Chandra G, Njie F, Berry J, Harvey A, et al. Pregnancy-related deaths: data from maternal mortality review committees in 36 US states, 2017-2019. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/reproductivehealth/maternal-mortality/erase-mm/data-mmrc.html [accessed 2023-03-21]
7. Ogunniyi MO, Mahmoud Z, Commodore-Mensah Y, Fleg JL, Fatade YA, Quesada O, et al. Eliminating disparities in cardiovascular disease for Black women: JACC review topic of the week. J Am Coll Cardiol 2022;80(18):1762-1771 [FREE Full text] [doi: 10.1016/j.jacc.2022.08.769] [Medline: 36302590]
8. Hoffman MK, Ma N, Roberts A. A machine learning algorithm for predicting maternal readmission for hypertensive disorders of pregnancy. Am J Obstet Gynecol MFM 2021;3(1):100250. [doi: 10.1016/j.ajogmf.2020.100250] [Medline: 33451620]

9. Yang L, Sun G, Wang A, Jiang H, Zhang S, Yang Y, et al. Predictive models of hypertensive disorders in pregnancy based on support vector machine algorithm. Technol Health Care 2020;28(S1):181-186 [FREE Full text] [doi: 10.3233/THC-209018] [Medline: 32364150]
10. Shara N, Anderson KM, Falah N, Ahmad MF, Tavazoei D, Hughes JM, et al. Early identification of maternal cardiovascular risk through sourcing and preparing electronic health record data: machine learning study. JMIR Med Inform 2022 Feb 10;10(2):e34932 [FREE Full text] [doi: 10.2196/34932] [Medline: 35142637]
11. Gil MM, Cuenca-Gómez D, Rolle V, Pertegal M, Díaz C, Revello R, et al. Validation of machine-learning model for first-trimester prediction of pre-eclampsia using cohort from PREVAL study. Ultrasound Obstet Gynecol 2024;63(1):68-74. [doi: 10.1002/uog.27478] [Medline: 37698356]
12. Ranjbar A, Taeidi E, Mehrnoush V, Roozbeh N, Darsareh F. Machine learning models for predicting pre-eclampsia: a systematic review protocol. BMJ Open 2023;13(9):e074705 [FREE Full text] [doi: 10.1136/bmjopen-2023-074705] [Medline: 37696628]
13. Garcia M, Mulvagh SL, Merz CNB, Buring JE, Manson JE. Cardiovascular disease in women: clinical perspectives. Circ Res 2016;118(8):1273-1293 [FREE Full text] [doi: 10.1161/CIRCRESAHA.116.307547] [Medline: 27081110]
14. Gad MM, Elgendy IY, Mahmoud AN, Saad AM, Isogai T, Mathias IS, et al. Disparities in cardiovascular disease outcomes among pregnant and post-partum women. J Am Heart Assoc 2021;10(1):e017832 [FREE Full text] [doi: 10.1161/JAHA.120.017832] [Medline: 33322915]
15. Davis SK, Gebreab S, Quarells R, Gibbons GH. Social determinants of cardiovascular health among Black and White women residing in stroke belt and buckle regions of the South. Ethn Dis 2014;24(2):133-143 [FREE Full text] [Medline: 24804357]
16. Chinn JJ, Martin IK, Redmond N. Health equity among Black women in the United States. J Womens Health (Larchmt) 2021;30(2):212-219 [FREE Full text] [doi: 10.1089/jwh.2020.8868] [Medline: 33237831]
17. Goddu AP, O'Conor KJ, Lanzkron S, Saheed MO, Saha S, Peek ME, et al. Do words matter? Stigmatizing language and the transmission of bias in the medical record. J Gen Intern Med 2018;33(5):685-691 [FREE Full text] [doi: 10.1007/s11606-017-4289-2] [Medline: 29374357]
18. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field. J Am Med Inform Assoc 2020;27(7):1132-1135 [FREE Full text] [doi: 10.1093/jamia/ocaa072] [Medline: 32324855]
19. Mehta LS, Warnes CA, Bradley E, Burton T, Economy K, Mehran R, et al. Cardiovascular considerations in caring for pregnant patients: a scientific statement from the American Heart Association. Circulation 2020;141(23):e884-e903 [FREE Full text] [doi: 10.1161/CIR.0000000000000772] [Medline: 32362133]
20. Magun E, DeFilippis EM, Noble S, LaSala A, Waksmonski C, D'Alton ME, et al. Cardiovascular care for pregnant women with cardiovascular disease. J Am Coll Cardiol 2020;76(18):2102-2113 [FREE Full text] [doi: 10.1016/j.jacc.2020.08.071] [Medline: 33121718]
21. Betts KS, Kisely S, Alati R. Predicting common maternal postpartum complications: leveraging health administrative data and machine learning. BJOG 2019;126(6):702-709. [doi: 10.1111/1471-0528.15607] [Medline: 30628159]
22. O'Neil S, Platt I, Vohra D, Pendl-Robinson E, Dehus E, Zephyrin LC, et al. The high costs of maternal morbidity show why we need greater investment in maternal health. The Commonwealth Fund. 2021. URL: https://www.commonwealthfund.org/publications/issue-briefs/2021/nov/high-costs-maternal-morbidity-need-investment-maternal-health [accessed 2024-03-23]
23. Roberts CT, Jankovic-Karasoulos T, Arthurs AL. Is the US failing women? EClinicalMedicine 2021;31:100701 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100701] [Medline: 33490929]
24. Ben-Israel D, Jacobs WB, Casha S, Lang S, Ryu WHA, de Lotbiniere-Bassett M, et al. The impact of machine learning on patient care: a systematic review. Artif Intell Med 2020;103:101785. [doi: 10.1016/j.artmed.2019.101785] [Medline: 32143792]
25. Dhombres F, Bonnard J, Bailly K, Maurice P, Papageorghiou AT, Jouannic JM. Contributions of artificial intelligence reported in obstetrics and gynecology journals: systematic review. J Med Internet Res 2022;24(4):e35465 [FREE Full text] [doi: 10.2196/35465] [Medline: 35297766]
26. Edvinsson C, Björnsson O, Erlandsson L, Hansson SR. Predicting intensive care need in women with preeclampsia using machine learning—a pilot study. Hypertens Pregnancy 2024;43(1):2312165 [FREE Full text] [doi: 10.1080/10641955.2024.2312165] [Medline: 38385188]
27. Alkhodari M, Xiong Z, Khandoker AH, Hadjileontiadis LJ, Leeson P, Lapidaire W. The role of artificial intelligence in hypertensive disorders of pregnancy: towards personalized healthcare. Expert Rev Cardiovasc Ther 2023;21(7):531-543. [doi: 10.1080/14779072.2023.2223978] [Medline: 37300317]
28. Leist AK, Klee M, Kim JH, Rehkopf DH, Bordas SPA, Muniz-Terrera G, et al. Mapping of machine learning approaches for description, prediction, and causal inference in the social and health sciences. Sci Adv 2022;8(42):eabk1942 [FREE Full text] [doi: 10.1126/sciadv.abk1942] [Medline: 36260666]

29. Bower JK, Patel S, Rudy JE, Felix AS. Addressing bias in electronic health record-based surveillance of cardiovascular disease risk: finding the signal through the noise. Curr Epidemiol Rep 2017;4(4):346-352 [FREE Full text] [doi: 10.1007/s40471-017-0130-z] [Medline: 31223556]
30. Hameed AB, Lawton ES, McCain CL, Morton CH, Mitchell C, Main EK, et al. Pregnancy-related cardiovascular deaths in California: beyond peripartum cardiomyopathy. Am J Obstet Gynecol 2015;213(3):379.e1-379.e10. [doi: 10.1016/j.ajog.2015.05.008] [Medline: 25979616]

Abbreviations

bpm: beats per minute
EHR: electronic health record
HELLP: hemolysis, elevated liver enzymes, and low platelet
HOPE-CAT: Healthy Outcomes for all Pregnancy Experiences-Cardiovascular-Risk Assessment Technology
MI: myocardial infarction
ML: machine learning

Edited by A Mavragani; submitted 29.09.23; peer-reviewed by J Chen, KM Kuo, F Albalwy; comments to author 16.01.24; revised version received 08.03.24; accepted 11.03.24; published 22.04.24.

Please cite as:
Shara N, Mirabal-Beltran R, Talmadge B, Falah N, Ahmad M, Dempers R, Crovatt S, Eisenberg S, Anderson K
Use of Machine Learning for Early Detection of Maternal Cardiovascular Conditions: Retrospective Study Using Electronic Health Record Data
JMIR Cardio 2024;8:e53091
URL: https://cardio.jmir.org/2024/1/e53091
doi:10.2196/53091
PMID:38648629

Original Paper

The Effect of an AI-Based, Autonomous, Digital Health Intervention Using Precise Lifestyle Guidance on Blood Pressure in Adults With Hypertension: Single-Arm Nonrandomized Trial

Jared Leitner[1], MS; Po-Han Chiang[1], PhD; Parag Agnihotri[2], MD; Sujit Dey[1], PhD

[1]Electrical and Computer Engineering Department, University of California, San Diego, La Jolla, CA, United States

[2]Department of Medicine, University of California, San Diego, La Jolla, CA, United States

Corresponding Author:
Jared Leitner, MS
Electrical and Computer Engineering Department
University of California, San Diego
9500 Gilman Dr
La Jolla, CA, 92093
United States
Phone: 1 8587220467
Email: jjleitne@ucsd.edu

Abstract

Background: Home blood pressure (BP) monitoring with lifestyle coaching is effective in managing hypertension and reducing cardiovascular risk. However, traditional manual lifestyle coaching models significantly limit availability due to high operating costs and personnel requirements. Furthermore, the lack of patient lifestyle monitoring and clinician time constraints can prevent personalized coaching on lifestyle modifications.

Objective: This study assesses the effectiveness of a fully digital, autonomous, and artificial intelligence (AI)–based lifestyle coaching program on achieving BP control among adults with hypertension.

Methods: Participants were enrolled in a single-arm nonrandomized trial in which they received a BP monitor and wearable activity tracker. Data were collected from these devices and a questionnaire mobile app, which were used to train personalized machine learning models that enabled precision lifestyle coaching delivered to participants via SMS text messaging and a mobile app. The primary outcomes included (1) the changes in systolic and diastolic BP from baseline to 12 and 24 weeks and (2) the percentage change of participants in the controlled, stage-1, and stage-2 hypertension categories from baseline to 12 and 24 weeks. Secondary outcomes included (1) the participant engagement rate as measured by data collection consistency and (2) the number of manual clinician outreaches.

Results: In total, 141 participants were monitored over 24 weeks. At 12 weeks, systolic and diastolic BP decreased by 5.6 mm Hg (95% CI −7.1 to −4.2; P<.001) and 3.8 mm Hg (95% CI −4.7 to −2.8; P<.001), respectively. Particularly, for participants starting with stage-2 hypertension, systolic and diastolic BP decreased by 9.6 mm Hg (95% CI −12.2 to −6.9; P<.001) and 5.7 mm Hg (95% CI −7.6 to −3.9; P<.001), respectively. At 24 weeks, systolic and diastolic BP decreased by 8.1 mm Hg (95% CI −10.1 to −6.1; P<.001) and 5.1 mm Hg (95% CI −6.2 to −3.9; P<.001), respectively. For participants starting with stage-2 hypertension, systolic and diastolic BP decreased by 14.2 mm Hg (95% CI −17.7 to −10.7; P<.001) and 8.1 mm Hg (95% CI −10.4 to −5.7; P<.001), respectively, at 24 weeks. The percentage of participants with controlled BP increased by 17.2% (22/128; P<.001) and 26.5% (27/102; P<.001) from baseline to 12 and 24 weeks, respectively. The percentage of participants with stage-2 hypertension decreased by 25% (32/128; P<.001) and 26.5% (27/102; P<.001) from baseline to 12 and 24 weeks, respectively. The average weekly participant engagement rate was 92% (SD 3.9%), and only 5.9% (6/102) of the participants required manual outreach over 24 weeks.

Conclusions: The study demonstrates the potential of fully digital, autonomous, and AI-based lifestyle coaching to achieve meaningful BP improvements and high engagement for patients with hypertension while substantially reducing clinician workloads.

Trial Registration: ClinicalTrials.gov NCT06337734; https://clinicaltrials.gov/study/NCT06337734

(JMIR Cardio 2024;8:e51916) doi:10.2196/51916

KEYWORDS

blood pressure; hypertension; digital health; lifestyle change; lifestyle medicine; wearables; remote patient monitoring; artificial intelligence; AI; mobile phone

Introduction

Background

High blood pressure (BP), or hypertension, is one of the most prevalent chronic diseases in the world [1]. Hypertension affects 48% (approximately 120 million) of adults in the United States, and 78% (approximately 93 million) of the cases are uncontrolled (ie, BP≥130/80 mm Hg) [2]. Hypertension is a major risk factor for stroke and acute myocardial infarction [3] and remains a large public health challenge with an extra cost of US $2000 per year per hypertension patient, resulting in an additional US $131 billion in annual health care costs in the United States [4]. The American College of Cardiology and American Heart Association's clinical practice guidelines define hypertension as systolic BP (SBP)≥130 mm Hg or diastolic BP (DBP)≥80 mm Hg, consistently over time [5]. A large-scale analysis of 48 randomized clinical trials showed that a 5–mm Hg reduction in SBP lowered the risk of major cardiovascular events by 10% [6], highlighting the importance of developing new strategies to achieve hypertension control at scale.

Hypertension management typically begins with home monitoring of BP to gain a more accurate estimate of a patient's BP within their usual, daily routine [7]. However, self-monitoring without additional support is not associated with lower BP or better control [8-10]. Lifestyle management in conjunction with self-monitoring is effective in controlling BP as lifestyle factors (eg, activity, sleep, diet, and stress) have a substantial impact on BP [11-14]. Even for patients taking antihypertensive medication, lifestyle management can enhance medication efficacy, leading to better BP control [15]. Traditionally, lifestyle management involves patients with hypertension visiting their primary care physician (PCP) and receiving guidance on lifestyle modifications that are generally known to improve BP. However, due to time constraints related to workload, physicians are often unable to optimally counsel patients on lifestyle modifications or personalize their guidance [16,17]. Due to insufficient guidance and the lack of feedback in between clinic visits, patients may implement some of these changes; however, patient engagement and compliance are generally suboptimal for achieving control. To improve patient engagement, new digital health technologies and remote patient monitoring programs have been developed for hypertension care [18-21]. These programs typically provide patients with remote monitoring devices (eg, BP cuffs and activity trackers) and match patients with health coaches. BP and lifestyle data collected from remote monitoring devices allow health coaches to view trends and make personalized recommendations to patients. However, these approaches do not consider the individual impact of lifestyle factors on BP, which may vary across individuals due to physiological differences. Furthermore, the reliance on health coaches is highly time and resource intensive, resulting in a high operating cost, which significantly limits scalability [22].

Objectives

To address the challenges of poor patient engagement due to generic, insufficient guidance and limited scalability of care due to human coaching models, we propose an artificial intelligence (AI)–driven, autonomous, precise lifestyle coaching program for patients with hypertension. The intervention platform consists of a monitoring system that ingests lifestyle and BP data and builds personalized machine learning (ML) models to determine the individual impact of different lifestyle factors on BP. On the basis of the lifestyle impact analysis, the system autonomously provides precise lifestyle recommendations delivered to a patient's smartphone that enable patients to focus on specific aspects of their lifestyle that have the greatest associations with their BP. While the platform autonomously engages patients, it is clinician supervised and notifies clinicians of critical BP readings. In our previous study [23], we enrolled 38 participants who were prehypertensive or had stage-1 hypertension (SBP between 120 and 139 mm Hg or DBP between 80 and 89 mm Hg) and demonstrated that 75% of the participants receiving the intervention were able to achieve a controlled BP (<130/80 mm Hg) after 16 weeks of engagement. However, the limitations of the previous study [23] are as follows: (1) the participants were not provided with an interactive mobile app for the delivery of our precise lifestyle recommendations, (2) the small number of participants did not enable rigorous evaluation, and (3) the study did not consider patients with stage-2 hypertension who can potentially benefit more from lifestyle management.

This study aims to evaluate the effectiveness of our AI-based, precise lifestyle guidance coaching program in helping patients with stage-2 hypertension achieve BP control and demonstrate the platform's scalability. The primary study objectives are to evaluate the change in BP and the percentage change of participants in different BP categories (controlled, stage-1 hypertension, and stage-2 hypertension) over time (baseline, 12 weeks, and 24 weeks). Secondary objectives include assessing participant engagement as measured by consistency of data collection and interactions with our mobile app and determining the number of manual clinician interventions, as defined by the escalation rules set for the study, to assess the potential scalability of our approach.

Methods

Recruitment

This study was performed in collaboration with the University of California, San Diego Health's Population Health Services Organization (PHSO). Participants were enrolled on a rolling basis from November 2021 to February 2023. The inclusion criteria required participants to have stage-2 hypertension (SBP≥140 mm Hg or DBP≥90 mm Hg per the American College of Cardiology and American Heart Association's 2017 guidelines [5]) based on their most recent clinical measurements and to be fully ambulatory (ie, not requiring an assistive device

such as a cane, wheelchair, or walker). In addition, participants were required to be aged ≥18 years at enrollment, be English speaking, and own an Android or iPhone (Apple Inc) smartphone. The trial was designed in a fully remote manner so that participants could participate entirely from home. The PHSO care team aggregated a list of patients who met the inclusion criteria and sent a recruitment flyer via bulk message using the Epic MyChart (Epic Systems Corporation) messenger. The flyer introduced the study and instructed patients to email the study team if they were interested in participating. After contacting the study team, eligible patients were asked to complete an electronic informed consent form. Patients who consented were sent a Fitbit Inspire 2 (Fitbit Inc) and a Bluetooth-enabled Omron Silver (Omron Corporation) BP monitor to collect their lifestyle and BP data for up to 6 months. Each shipment included instructions for self-onboarding, which described the steps to set up and connect the devices to the patient's mobile phone. Patients who already owned a Fitbit or Apple Watch (Apple Inc) had the option to use their device instead of receiving one from the study team. Patients who required an extra-large cuff were provided an iHealth Ease (iHealth Labs Inc) BP monitor instead of an Omron Silver.

Ethical Considerations

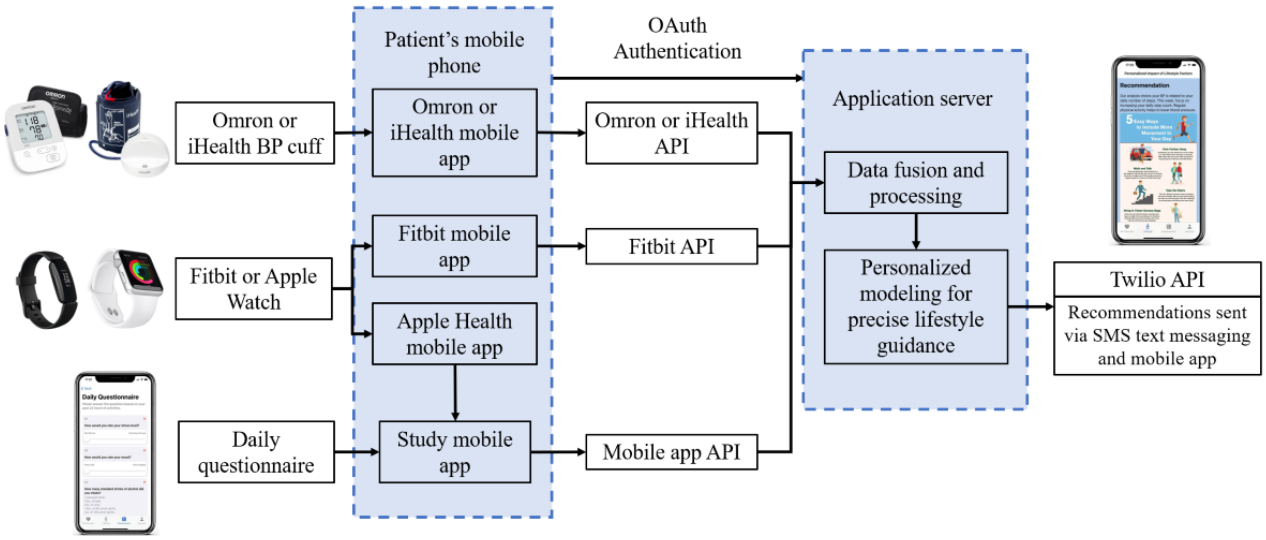
This study (protocol #181405) was reviewed and approved by the University of California, San Diego's Human Research Protections Program, which operates Institutional Review Boards. All participants in this study provided informed consent, which included the collection of their data and the provision of study results derived from their individual data. The confidentiality and privacy of participants were ensured by assigning a deidentified code to each patient. While participants were not offered monetary compensation, those without a BP monitor or wearable device were provided with these devices. The study was registered at ClinicalTrials.gov (NCT06337734).

Study Design and Data Collection

We collected data from each participant using a Fitbit or Apple Watch, Omron or iHealth wireless BP monitor, and the study's questionnaire mobile app. Participants were asked to wear their Fitbit or Apple Watch as often as possible, including during sleep, and take 1 to 2 BP measurements per day, in the morning (8 AM-10 AM) or evening (7 PM-9 PM). We provided participants with instructions on how to take accurate resting BP readings [24] and asked that they take 3 consecutive readings during each morning and evening session. This resulted in 1 to 2 sets of 3 measurements per day, and the average of the 3 measurements was used as the final value for each session. Participants synced their BP data to the Omron or iHealth mobile app and their Fitbit data to the Fitbit mobile app; subsequently, the data were automatically uploaded to the Omron, iHealth, or Fitbit clouds. These data were retrieved remotely through the application programming interfaces (APIs) provided by Omron, iHealth, and Fitbit. Data from the Apple Watch were synced with the study mobile app and uploaded via a custom API to our server. In addition, participants completed a daily questionnaire using our study mobile app that asked about their stress, mood, and dietary choices over the past 24 hours. These questions were developed in collaboration with physicians on our team. The diet questions are tailored to measure information relevant to hypertension, including alcohol, red meat, fruits or vegetables, and salt consumption [25]. The details of the questionnaire are described in our previous study [23]. In addition, we asked participants to complete a study experience survey that asked them to rate the difficulty level of completing the study tasks, how useful they found the recommendations, and their experience using the app. These responses were collected through the mobile app and used to assess participant experience. Figure 1 describes the system architecture and data transmission.

Wrist-worn activity and sleep trackers have been widely used in health-related research studies [26], and devices such as Fitbits and Apple Watches have been shown to accurately measure parameters such as step count, heart rate, and sleep duration [27,28]. Fitbits and Apple Watches include an optical heart rate monitor and a 3-axis accelerometer. The devices use these sensors to calculate various health parameters, including lifestyle and vitals measurements. Lifestyle factors include activity (eg, steps, walking and running speed, and active time), sleep timing (eg, sleep duration, bedtime, and uptime), and sleep stages (ie, deep, light, rapid eye movement, and awake). These lifestyle factors are used as part of the intervention, in which we use ML techniques to determine which of the factors have the greatest association with a participant's BP and base our guidance on this analysis.

Figure 1. Architecture of data transmission. Participant data were collected from Bluetooth-enabled blood pressure (BP) monitors, wearable devices, and a mobile app–based questionnaire. Data were uploaded through the respective application programming interfaces (APIs) to our app server, where the individualized analysis was carried out before delivering recommendations to participants.



Description of the Intervention

The intervention is intended to support participants' daily efforts to improve BP and overall cardiometabolic function by facilitating behavioral changes that target physical activity, sleep hygiene, stress management, and dietary choices most relevant to their BP. The intervention platform uses remotely collected lifestyle and BP data to provide personalized, precise, and proactive lifestyle coaching using AI to participants with hypertension. The system integrates the data described in the previous section into a combined data set for each participant. Each participant's personal data set consists of lifestyle features (eg, step count, sleep duration, and salt consumption) that are time aligned with their BP measurements, which serve as the labels for training the ML model. Therefore, each participant's data set is used to train a personal ML model that can predict BP using the participant's lifestyle data as input. With this trained model, the intervention system can determine how different aspects of lifestyle affect the participant's BP. On the basis of the model's determination of the lifestyle factors' impact, the system generates precise lifestyle recommendations. Each lifestyle factor is mapped to a corresponding lifestyle recommendation that was designed with physicians on our team to be consistent with evidence-based clinical guidelines. Furthermore, prior studies have demonstrated that these recommendations, such as increasing step count [29,30], improving sleep quality [31,32], managing stress [33], and reducing salt consumption [34,35], can result in BP reduction. The objective of these precise lifestyle recommendations is to encourage participants to concentrate on 1 aspect of their lifestyle at a time, focusing on the factor that has the greatest association with their BP based on the underlying relationship between their BP and lifestyle factors. We describe the AI-based intervention platform in more detail in our previous study [23].

Participants received weekly lifestyle recommendations based on their data and personalized analytics, which continuously evolved over time. These recommendations were delivered to participants via programmable text messages using the Twilio API (Twilio Inc) service [36] and were displayed in the study mobile app. Each text message included a summary of the participant's BP progression for the current week in addition to the lifestyle recommendation. Figure 2 displays examples of these weekly lifestyle recommendations provided in the study app. In addition, patients completed a midweek check-in on the app, which asked whether they could follow each recommendation (yes or no) and to rate the recommendation difficulty on a scale from 1 to 5.

The system includes a safety mechanism to involve clinician intervention in the case of critically high or low BP readings. Critically high BP was defined as SBP>180 mm Hg or DBP>110 mm Hg, and critically low BP was defined as SBP<90 mm Hg or DBP<60 mm Hg [5]. After a critical reading, participants received a text message asking them to remeasure their BP and prompting them to seek assistance or call their medical provider if they were experiencing certain symptoms (eg, chest pain and severe headache). After 2 critical readings in a row, an escalation notification was sent to the PHSO care team via email for manual outreach. To avoid notification fatigue, we limited the number of critically high or low BP notifications sent to the care team to 1 notification per week for a patient.

Figure 2. Lifestyle recommendations delivered in the mobile app. Participants received weekly lifestyle recommendations based on their data and personalized analytics. The recommendations encouraged participants to prioritize a single lifestyle modification at a time, focusing on the factor that had the greatest impact on their blood pressure (BP).

Primary Outcomes: BP Change and Population Hypertension Control

The first primary outcome was the change in SBP and DBP from baseline to 12 weeks and 24 weeks. A participant's baseline BP was calculated as the average of their readings during the first week of the study. The 12th- and 24th-week BPs were a participant's average reading during that week of the study plus 1 week and minus 1 week. We included BP measurements from 1 week before and after to get a more representative result. For example, the 12-week value was the average of all readings from weeks 11 to 13. As previously mentioned, a 5–mm Hg reduction in SBP can lower the risk of major cardiovascular events by 10% [6]. This motivated us to determine the percentage of participants who experienced >5–mm Hg reduction in SBP at 12 weeks and 24 weeks. To understand the effect on participants with different baseline BPs, we carried out subgroup analysis in which participants were sorted into 3 groups based on their baseline BP: (1) controlled (SBP<130 mm Hg and DBP<80 mm Hg), (2) stage-1 hypertension (SBP 130-139 mm Hg or DBP 80-89 mm Hg), and (3) stage-2 hypertension (SBP≥140 mm Hg or DBP≥90 mm Hg).

Another primary outcome was the percentage change of participants in different BP categories from baseline to 12 weeks and 24 weeks. To assess this, we calculated the percentage of participants who were in the controlled, stage-1 hypertension, and stage-2 hypertension categories at baseline, 12 weeks, and 24 weeks. Using these percentages, we determined the percentage change from baseline to 12 weeks and 24 weeks.

Secondary Outcomes: Participant Engagement and Clinician Intervention

A secondary outcome measured participant engagement as determined by the consistency of data collection and interactions with our mobile app. The 3 main tasks participants were asked to complete included measuring BP, syncing their wearable device, and answering the mobile app questionnaire. As a result, we used these 3 tasks as our measure of engagement and calculated the percentage of participants completing each of these tasks each week. A participant was marked as engaged for a given week if they provided a BP reading, synced their wearable device data, and answered the questionnaire at least once during the week.

Another secondary outcome was the number of times participants were escalated to the PHSO care team for manual follow-up. The objective of this outcome was to determine the care team's time and resource requirements to implement the intervention and assess the scalability of our approach. The condition for care team intervention was 2 critical BP readings in a row, as previously described.

Statistical Analysis

Descriptive statistics (eg, mean, SD, and percentage) were calculated to describe the demographic and baseline clinical characteristics of the enrolled study population. We compared the characteristics between subgroups based on their baseline BP classification.

Change in SBP and DBP from baseline to 12 weeks and 24 weeks was analyzed using a 2-tailed paired Student *t* test with the level of statistical significance set to $P<.05$. Furthermore, 95% CIs were calculated for these changes. Baseline and follow-up BP data were normally distributed. The McNemar nonparametric test was used to examine the change in the proportion of participants in the controlled, stage-1, and stage-2 BP range from baseline to 12 weeks and 24 weeks. The McNemar test is used to determine if there is a statistically significant difference in proportions between paired data. We conducted all statistical analyses with Python 3.9 (Python Software Foundation) using the *NumPy*, *Pandas*, and *SciPy* libraries.

Results

Feasibility Outcomes: Recruitment, Adherence, and Participant Experience

Participants were enrolled on a rolling basis from November 2021 to February 2023. Figure 3 details the recruitment numbers and participant flow through the study. A total of 274 patients responded to the Epic MyChart recruitment message by contacting our team and expressing interest. In total, 164 patients consented to join the study, out of which 141 (86%) were onboarded and started collecting data. There was a 9.2% (13/141) dropout rate from the start of the study to 12 weeks and a 20.3% (26/128) dropout rate from 12 weeks to 24 weeks. Reasons for participants withdrawing from the study included receiving new medical diagnoses (eg, cancer diagnosis), achieving a healthy BP, family emergencies, and other personal reasons. For the 141 participants who onboarded, Table 1 compares the characteristics between subgroups based on baseline BP classifications. The average age of participants was 57.5 (SD 13.9) years, and 44% (62/141) of the participants were female. For participants who had stage 2 hypertension at baseline, the average baseline BP was 141.9/89.4 mm Hg. In total, 83.7% (118/141) of the participants reported that they were taking antihypertensive medication at the beginning of the study.

As previously described, we asked participants each week to rate the difficulty of the recommendations they received on a scale from 1 to 5 and indicate whether they could follow each recommendation. This was done to assess compliance and the perceived difficulty of the recommendations. The histogram of difficulty ratings, divided into *Yes* and *No* responses, is shown in Multimedia Appendix 1. Recommendations were followed 63.64% (721/1133) of the time and not followed 36.36% (412/1133) of the time. The average difficulty rating for recommendations that were followed was 1.97, indicating lower difficulty, whereas the average for those not followed was 3.67, indicating higher difficulty. Evidently, there is a negative correlation between the perceived difficulty of a recommendation and its likelihood of being followed. We also tracked the number of unique recommendations each patient was sent. Out of the 37 unique recommendations, patients received an average of 9.4 (25%) unique recommendations each. The distribution of the number of unique recommendations is shown in Figure 4. The median and IQR suggest a distribution close to normal. The maximum number of unique recommendations received by a single patient was as high as 21. These statistics demonstrate a broad range of recommendations given to the patients, covering various aspects of lifestyle.

An additional feasibility outcome we evaluated was participant experience as measured by responses to a study experience survey. As previously mentioned, this survey asked patients to rate the difficulty level of completing the study tasks, how useful they found the recommendations, and their experience using the app. Multimedia Appendix 2 presents the distribution of participant responses to these 3 questions. In total, 70 participants responded to the survey. In total, 61% (43/70) of the participants responded that the study tasks were "easy" or "very easy" to incorporate into their daily routine, 51% (36/70) of the participants found the personalized recommendations to be "useful" or "very useful" compared to generic recommendations, and 86% (60/70) of the participants rated the app experience as "good" or "great."

Figure 3. Flow of participants through the study. Adults with hypertension were enrolled from the University of California, San Diego Health between November 2021 and February 2023 into a single-arm nonrandomized trial. BP: blood pressure.

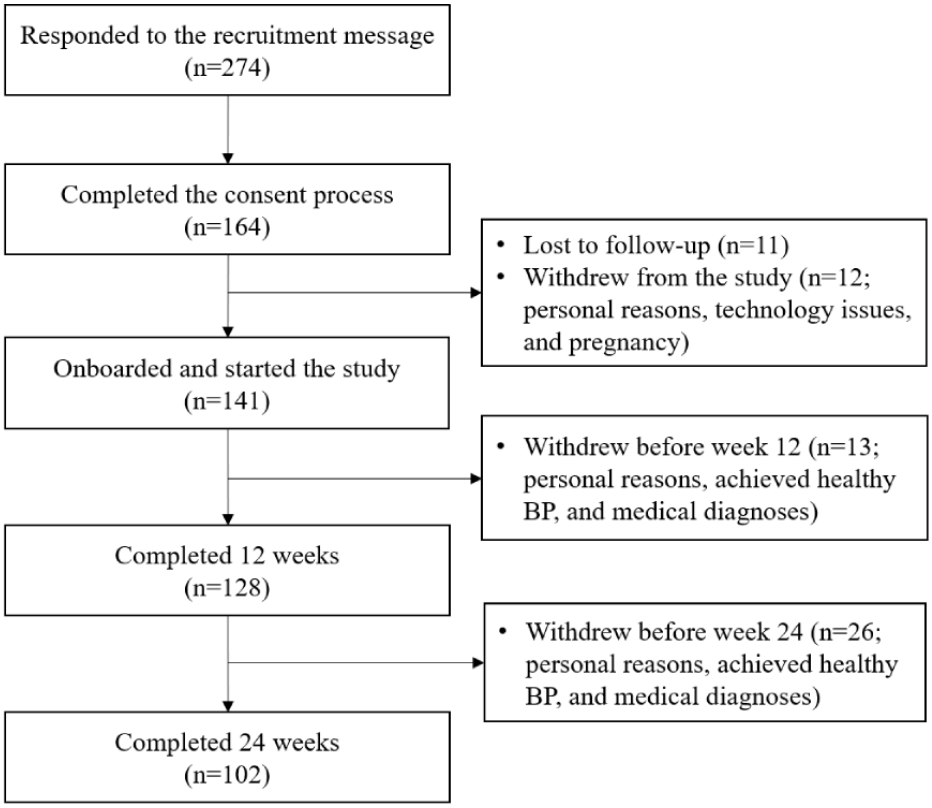


Table 1. Participant demographics and characteristics grouped by baseline BP[a] (N=141).

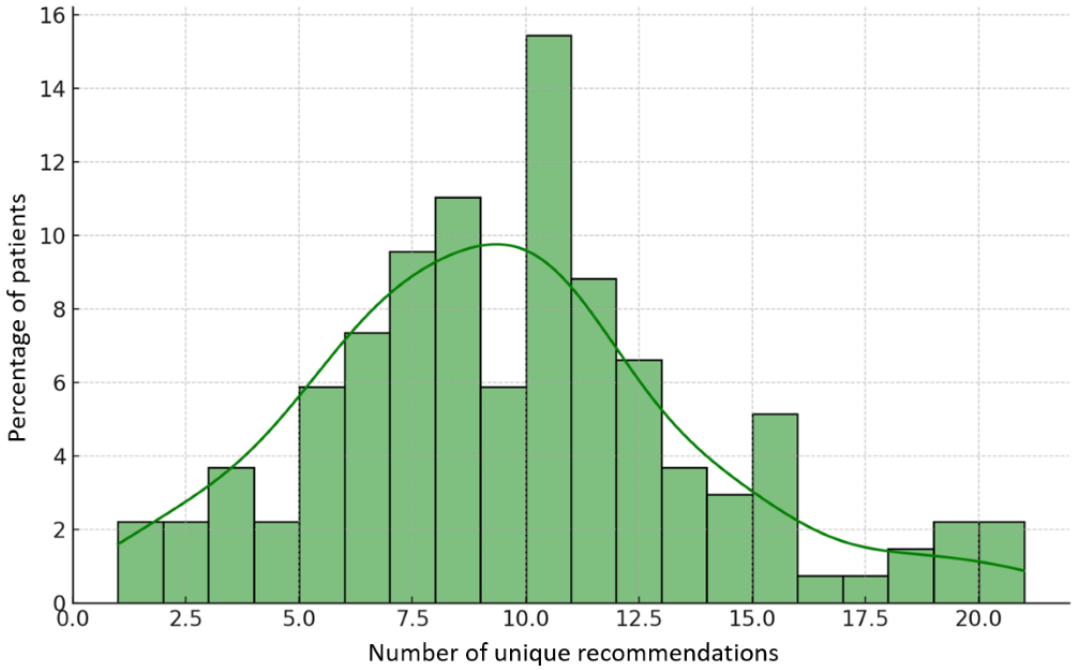
Characteristics	Baseline BP category			
	All (N=141)	Controlled (n=38)	Stage 1 (n=48)	Stage 2 (n=55)
Age (y), mean (SD)	57.5 (13.9)	57.8 (16.0)	57.6 (12.6)	57.3 (13.5)
Female, n (%)	62 (44)	14 (37)	24 (50)	24 (44)
Weight (lb), mean (SD)	175.8 (48.4)	170.0 (41.6)	164.5 (52.3)	189.7 (45.7)
Baseline SBP[b] (mm Hg), mean (SD)	131.9 (11.5)	121.4 (6.1)	128.8 (7.1)	141.9 (9.3)
Baseline DBP[c] (mm Hg), mean (SD)	82.9 (9.0)	74.2 (4.4)	82.2 (6.4)	89.4 (8.0)
Taking hypertension medication, n (%)	118 (83.7)	32 (84)	39 (81)	47 (85)

[a]BP: blood pressure.

[b]SBP: systolic blood pressure.

[c]DBP: diastolic blood pressure.

Figure 4. Distribution showing the number of unique recommendations sent to each patient. Patients received an average of 9.4 unique recommendations each.



BP Outcomes

For assessing BP outcomes, we used data from the 128 and 102 participants who completed 12 and 24 weeks in the study, respectively. Table 2 details the change in BP from baseline to 12 weeks. Across all participants, there was a statistically significant change of −5.6 mm Hg (95% CI −7.1 to −4.2; t_{127}=7.6; P<.001) in SBP and −3.8 mm Hg (95% CI −4.7 to −2.8; t_{127}=7.7; P<.001) in DBP after 12 weeks. Notably, 45.3% (58/128) of the participants achieved a clinically meaningful SBP drop of ≥5 mm Hg after 12 weeks. Table 3 details the change in BP from baseline to 24 weeks. For the participants who completed 24 weeks in the study, there was a statistically significant change of −8.1 mm Hg (95% CI −10.1 to −6.1; t_{101}=8.1; P<.001) in SBP and −5.1 mm Hg (95% CI −6.2 to −3.9; t_{101}=8.4; P<.001) in DBP. In total, 58.8% (60/102) of the participants achieved a clinically meaningful SBP drop of ≥5 mm Hg after 24 weeks.

Table 2. Comparison of average BP[a] change at 12 weeks for different participant subgroups based on baseline BP (n=128)[b].

BP and subgroup	Participants, n (%)	Change in BP at 12 weeks, Δmean (SD; 95% CI)	*t* test (*df*)	*P* value	≥5–mm Hg reduction in SBP[c] at 12 weeks, n (%)
SBP					
Overall	128 (100)	−5.6 (8.1; −7.1 to −4.2)	7.6 (127)	<.001	58 (45.3)
Controlled	31 (24.2)	−3.6 (5.2; −5.5 to −1.6)	3.7 (30)	.001	11 (35)
Stage 1	46 (35.9)	−2.6 (7.2; −4.8 to −0.5)	2.5 (45)	.02	14 (30)
Stage 2	51 (39.8)	−9.6 (9.2; −12.2 to −6.9)	7.3 (50)	<.001	33 (65)
DBP[d]					
Overall	128 (100)	−3.8 (5.5; −4.7 to −2.8)	7.7 (127)	<.001	N/A[e]
Controlled	31 (24.2)	−1.6 (3.8; −3.0 to −0.2)	2.3 (30)	.03	N/A
Stage 1	46 (35.9)	−3.1 (4.4; −4.4 to −1.7)	4.7 (45)	<.001	N/A
Stage 2	51 (39.8)	−5.7 (6.7; −7.6 to −3.9)	6.2 (50)	<.001	N/A

[a]BP: blood pressure.

[b]For participants with stage-2 hypertension at baseline, SBP and DBP changed by −9.6 mm Hg and −5.7 mm Hg, respectively, after 12 weeks.

[c]SBP: systolic blood pressure.

[d]DBP: diastolic blood pressure.

[e]N/A: not applicable.

Table 3. Comparison of average BP[a] change at 24 weeks for different participant subgroups based on baseline BP (n=102)[b].

BP and subgroups	Participants, n (%)	Change in BP at 24 weeks, Δmean (SD; 95% CI)	*t* test (*df*)	*P* value	≥5–mm Hg reduction in SBP[c] at 24 weeks, n (%)
SBP					
Overall	102 (100)	−8.1 (10.1; −10.1 to −6.1)	8.1 (101)	<.001	60 (58.8)
Controlled	28 (27.5)	−3.9 (8.6; −7.1 to −0.8)	2.6 (27)	.02	14 (50)
Stage 1	37 (36.3)	−5.2 (8.0; −7.9 to −2.5)	3.9 (36)	<.001	17 (46)
Stage 2	37 (36.3)	−14.2 (10.6; −17.7 to −10.7)	8.2 (36)	<.001	29 (78)
DBP[d]					
Overall	102 (100)	−5.1 (6.0; −6.2 to −3.9)	8.4 (101)	<.001	N/A[e]
Controlled	28 (27.5)	−1.9 (4.3; −3.6 to −0.2)	2.3 (27)	.03	N/A
Stage 1	37 (36.3)	−4.4 (4.7; −6.0 to −2.8)	5.7 (36)	<.001	N/A
Stage 2	37 (36.3)	−8.1 (6.9; −10.4 to −5.7)	7.0 (36)	<.001	N/A

[a]BP: blood pressure.

[b]For participants with stage-2 hypertension at baseline, SBP and DBP changed by −14.2 mm Hg and −8.1 mm Hg, respectively, after 24 weeks.

[c]SBP: systolic blood pressure.

[d]DBP: diastolic blood pressure.

[e]N/A: not applicable.

Participants with a baseline BP classified as stage-2 hypertension had the greatest change in BP and the greatest percentage of participants achieving a clinically meaningful SBP drop after 12 and 24 weeks. For these participants, SBP and DBP improved by −9.6 mm Hg (95% CI −12.2 to −6.9; t_{50}=7.3; *P*<.001) and −5.7 mm Hg (95% CI −7.6 to −3.9; t_{50}=6.2; *P*<.001) after 12 weeks, respectively, and −14.2 mm Hg (95% CI −17.7 to −10.7; t_{36}=8.2; *P*<.001) and −8.1 mm Hg (95% CI −10.4 to −5.7; t_{36}=7.0; *P*<.001) after 24 weeks, respectively. In total, 65% (33/51) and 78% (29/37) of the participants achieved a clinically meaningful SBP drop of ≥5 mm Hg after 12 and 24 weeks, respectively.

Another primary outcome we assessed was the percentage change of participants in different BP categories from baseline to 12 weeks and 24 weeks. Tables 4 and 5 detail this analysis. For participants completing 12 weeks in the study, the percentage of participants in the controlled range increased by 17.2% from 24.2% (31/128) to 41.4% (53/128; McNemar χ^2_1=3.0, *P*<.001). The percentage of participants with stage 2 hypertension decreased by 25% from 39.8% (51/128) to 14.8% (19/128; McNemar χ^2_1=4.0, *P*<.001) after 12 weeks. This means that 63% (32/51) of the patients with stage-2 hypertension at baseline moved into lower BP categories after 12 weeks. For those who completed 24 weeks in the study, the percentage in the controlled range increased by 26.5% from 27.5% (28/102) to 53.9% (55/102; McNemar χ^2_1=2.0, *P*<.001), and the stage-2 percentage decreased by 26.5% from 36.3% (37/102) to 9.8% (10/102; McNemar χ^2_1=3.0, *P*<.001). This means that 73% (27/37) of the patients with stage-2 hypertension at baseline moved into lower BP categories after 24 weeks. Note that the percentage changes for the stage-1 hypertension category from baseline to 12 weeks and 24 weeks were not statistically significant at the *P*=.05 level. The smaller change in the stage-1 hypertension population is due to a cascading effect where the number of participants moving from stage 2 into stage 1 was offset by the number of patients moving out of stage 1 and into the controlled BP category. For example, from baseline to 24 weeks, 18 participants moved from stage 2 to stage 1, and 17 participants moved from stage 1 to the controlled category.

Table 4. Change in the percentage of participants in different BP[a] categories from baseline to 12 weeks (n=128)[b].

Subgroups	Population at baseline, n (%)	Population at 12 weeks, n (%)	12-week difference, n (%)	McNemar χ^2 (*df*)	*P* value
Controlled	31 (24.2)	53 (41.4)	22 (17.2)	3.0 (1)	<.001
Stage 1	46 (35.9)	56 (43.8)	10 (7.8)	20.0 (1)	.20
Stage 2	51 (39.8)	19 (14.8)	−32 (−25)	4.0 (1)	<.001

[a]BP: blood pressure.

[b]The percentage of participants with stage-2 hypertension decreased by 25% from 39.8% to 14.8% after 12 weeks.

Table 5. Change in the percentage of participants in different BP[a] categories from baseline to 24 weeks (n=102)[b].

Subgroups	Population at baseline, n (%)	Population at 24 weeks, n (%)	24-week difference, n (%)	McNemar χ^2 (*df*)	*P* value
Controlled	28 (27.5)	55 (53.9)	27 (26.5)	2.0 (1)	<.001
Stage 1	37 (36.3)	37 (36.3)	0 (0)	N/A[c]	N/A
Stage 2	37 (36.3)	10 (9.8)	−27 (−26.5)	3.0 (1)	<.001

[a]BP: blood pressure.

[b]The percentage of participants with stage-2 hypertension decreased by 26.5% from 36.3% to 9.8% after 24 weeks.

[c]N/A: not applicable.

Participant Engagement

We assessed participant engagement based on the percentage of active participants completing the program tasks each week. Figures 5-7 show the weekly percentage of active patients measuring their BP, syncing their wearable device, and answering the questionnaire during the 24 weeks, respectively. We set an engagement goal of 90% for the study, which is represented by the red dashed lines in the figures. The average BP measurement engagement rate was 93% (SD 4.3%), and this rate was >90% for 19 (79%) out of 24 weeks. The average wearable syncing engagement rate was 94% (SD 2.4%), and this rate was >90% for 21 (88%) out of 24 weeks. The average questionnaire engagement rate was 88% (SD 4.9%), and this rate was >90% for 10 (42%) out of 24 weeks.

Figure 5. Percentage of active participants measuring their blood pressure (BP) during the 24 weeks.

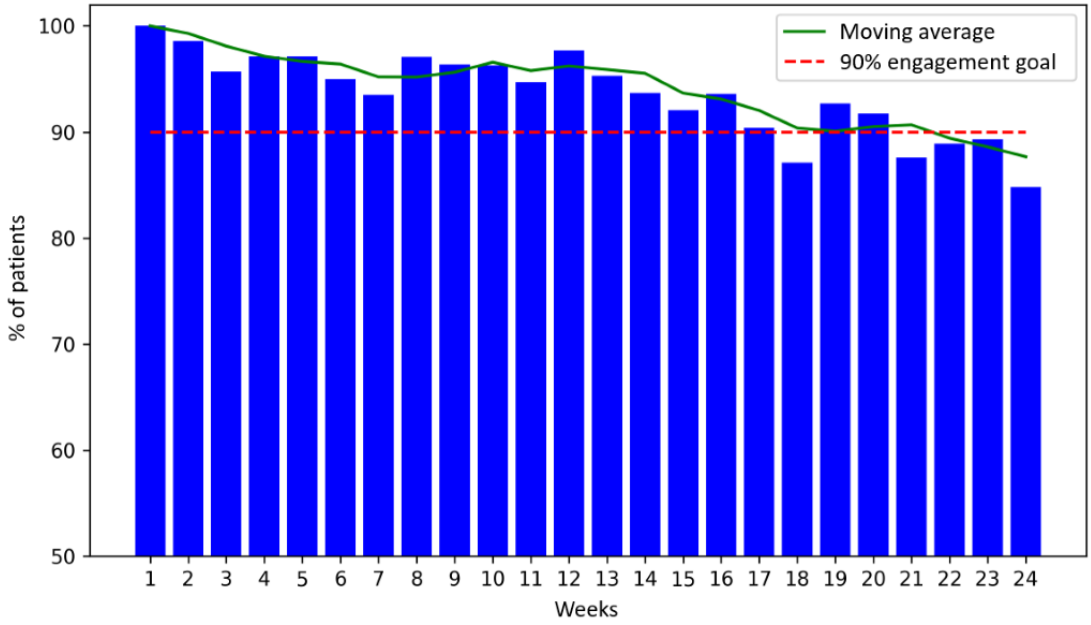


Figure 6. Percentage of active participants syncing their wearable device during the 24 weeks.

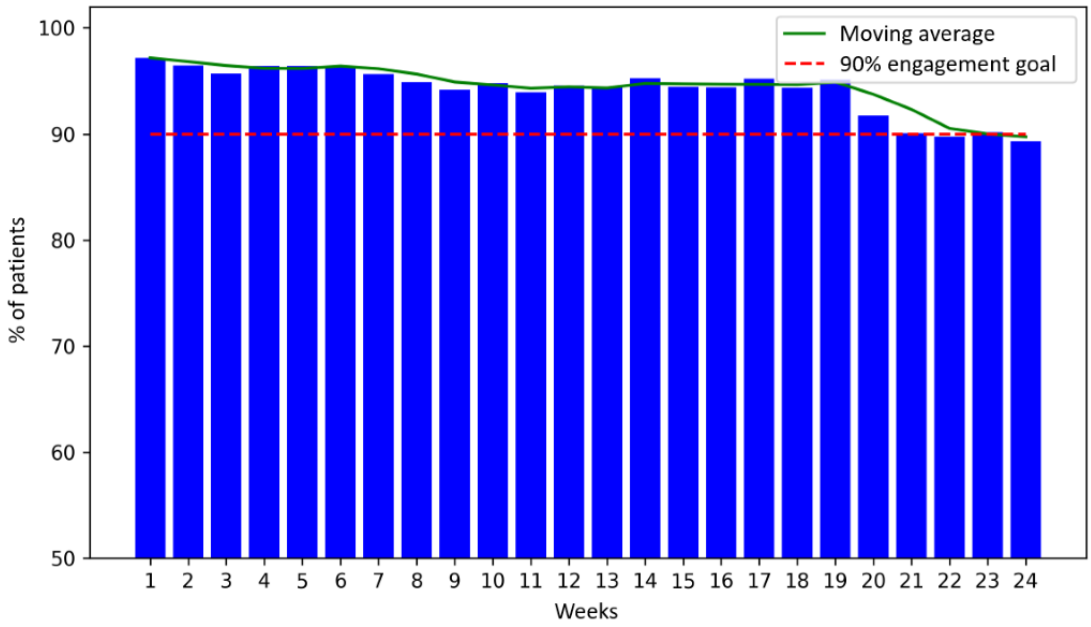
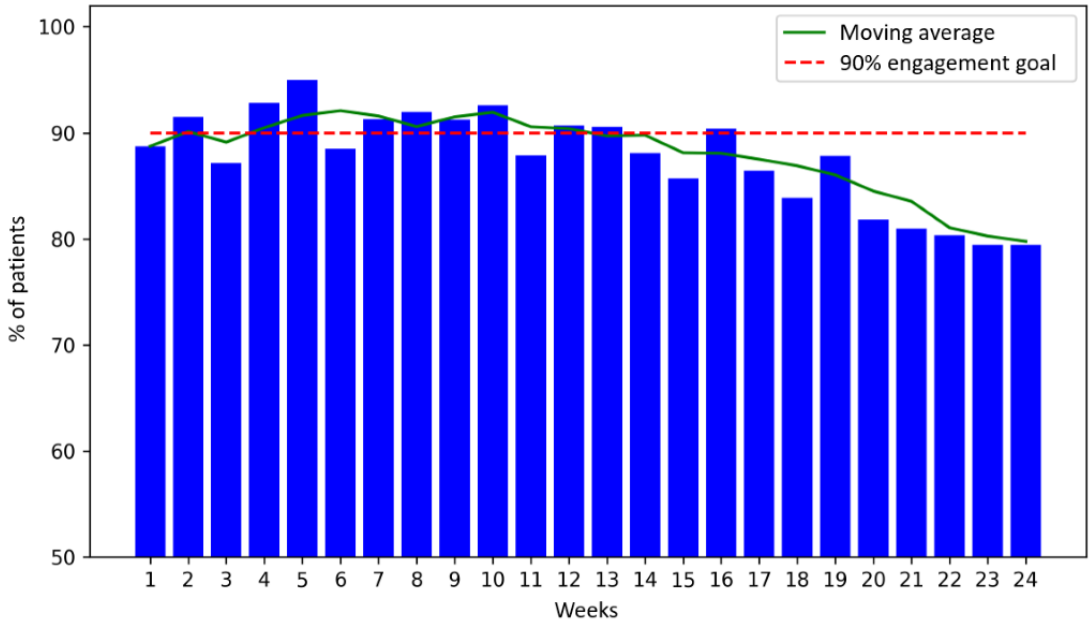


Figure 7. Percentage of active participants answering the questionnaire during the 24 weeks.



Clinician Intervention

For the 128 participants completing 12 weeks in the study, an escalation notification was sent to the care team 8 times. There were 3.9% (5/128) unique patients who required manual outreach during the first 12 weeks. For the 102 patients completing 24 weeks in the study, an escalation notification was sent to the PHSO care team 11 times. There were 5.9% (6/102) unique patients who required manual outreach during the 24 weeks.

Discussion

Principal Findings

This study aims to assess the effectiveness of a fully digital, autonomous, and AI-based lifestyle coaching program in achieving BP control and high engagement among adults with hypertension. The key components of this program included detailed lifestyle data collection via both wearables and questionnaires and weekly lifestyle recommendations based on personalized, AI-based analytics delivered via a mobile app. The guidance supported the participant's daily efforts to improve BP through behavioral changes that targeted physical activity, sleep hygiene, stress management, and dietary choices. Specifically, the program provided weekly guidance based on associations between lifestyle data and BP uncovered using ML and asked the participants to focus on the lifestyle factor with the greatest association. The precise lifestyle recommendations enabled participants to focus on the most relevant aspect of their lifestyle as opposed to receiving general guidance. Our intervention approach aligns with the Fogg Behavioral Model, which states that 3 elements (ability, motivation, and prompts)

are essential for behavior change [37]. By directing participants to focus on 1 lifestyle behavior at a time, the intervention simplified compliance and therefore increased the ability of the participants to adhere to the recommendations. This targeted strategy likely bolstered participants' motivation, as they could clearly see how specific lifestyle modifications directly influenced their BP. Each recommendation was delivered via a text message and prompted the user to take specific action. Furthermore, each recommendation was sent with a motivational message regarding their BP progress. We believe that this combination of personalized advice, ease of compliance, and motivational reinforcement contributed to our high engagement and improved BP outcomes.

We assessed multiple feasibility outcomes, including enrollment rate, adherence, and participant experience. In total, 59.9% (164/274) of the patients who initially expressed interest in joining the program ended up enrolling. Furthermore, although patients were recruited based on their last clinical BP reading, which required an SBP≥140 mm Hg or DBP≥90 mm Hg (stage-2 hypertension), many participants were not in the stage-2 range at baseline. Possible reasons for this include white coat hypertension [38] or that between the time of their last clinical BP reading and their enrollment in the study, they may have started taking BP medication or changed their diet. To improve the enrollment rate and ensure that patients who enroll have stage-2 hypertension, a new recruitment strategy is required. This new strategy could involve recruiting patients through PCP referrals. We hypothesize that this will increase the take-up rate due to increased trust from the more personal nature of the referral [39]. Furthermore, for the patients who are referred to the study, their PCPs would be instructed not to start the patients on any new BP medication or lifestyle intervention before the study, except in critical cases. This would help ensure patients joining the study are indeed in the stage-2 hypertension category. Another feasibility outcome we assessed was participant experience. While most participants (43/70, 61%) found the study tasks easy to incorporate into their daily routine, a few (3/70, 4%) found it difficult. These included difficulty in measuring BP due to work schedules and travel, caregiving responsibilities, and equipment and syncing issues. To address these challenges, the intervention should be more context aware and adapt the program tasks and recommendations based on patients' circumstances. For example, a patient who works a night shift should not be asked to measure their BP at the same time or be given the same sleep recommendations as a patient who works during the day. Context-aware interventions would enhance the patient experience and increase the engagement rate.

Participants experienced a statistically significant decrease of 8.1 mm Hg and 5.1 mm Hg in SBP and DBP, respectively, after 24 weeks. Furthermore, this improvement was more pronounced in participants who started the program with stage-2 hypertension, achieving a 14.2 mm Hg and 8.1 mm Hg reduction in SBP and DBP, respectively. Reducing BP holds clinical significance not only for individuals with stage 2 hypertension but also for those with elevated BP or stage 1 hypertension. This is clinically meaningful as lower SBP values have been associated with progressively reduced risks of stroke, major cardiovascular events, and cardiovascular as well as all-cause mortalities [40]. In addition to BP improvement, the study demonstrates the intervention's ability to maintain sustained engagement. However, the engagement rate dropped during the last 4 weeks potentially because the participants whose BP had improved through the program may have reduced their engagement as they did not feel the urgent need. In this study, the participant tasks remain consistent; however, participants may find it useful if the requirements are adaptive based on their health condition and preferences. It is worthwhile to design a dynamic mechanism that can adjust the extent and frequency of patient requirements based on the intervention progress. Both the BP and engagement results are achieved with minimal clinician intervention, primarily due to the autonomous nature of the intervention, demonstrating the potential scalability of this approach for hypertension management.

The observed BP improvement results from this study are comparable to those from clinician-led hypertension management programs [18-21]. The 3-month intervention program presented in the study by Wilson-Anumudu et al [18] combined lifestyle counseling with hypertension education, guided home BP monitoring, and support for taking medications and was led by either a registered nurse or certified diabetes care and education specialist. Patients with stage-2 hypertension who participated in this program experienced a 10.3 mm Hg and 6.5 mm Hg reduction in SBP and DBP, respectively, after 3 months. In the study by Milani et al [20], the 3-month digital intervention involved patients measuring their BP at least once per week and corresponding with pharmacists and health coaches to cocreate their treatment plan by choosing among various lifestyle modifications (eg, reducing dietary sodium) and medication options (eg, switching to generics or lower cost options). Patients with stage-2 hypertension participating in this program experienced a 14.0 mm Hg and 5.0 mm Hg reduction in SBP and DBP, respectively, after 3 months. Both interventions presented in the studies by Wilson-Anumudu et al [18] and Milani et al [20] assigned participants a designated hypertension coach who would provide lifestyle education and recommendations. These previous studies [18,20] primarily attribute their BP outcomes to the program's support led by health professionals who interpreted BP data and supported lifestyle change. While health coach–based programs can produce meaningful BP improvements, the reliance on health coaches is highly time and resource intensive. Consequently, these approaches have limited scalability and accessibility as an individual health coach can only engage and care for a limited number of patients at a time. In contrast, our results demonstrate that a fully digital, AI-based lifestyle coaching program can produce clinically meaningful BP improvements comparable to those of programs led by health professionals. There is also potential for our approach to be used in conjunction with health coach–based programs. Under such a framework, our AI-based interactions and learnings from the patients can extend the reach of health coaches and provide them with more detailed insights about lifestyle factors impacting patients.

Study Limitations and Future Directions

As this was a single-arm nonrandomized study, it was not possible to conduct a causal analysis due to the lack of a control

group. In addition, regression to the mean is another limitation as participants with initially high BP values may naturally converge toward the average over time. Therefore, to conduct causal analysis and account for regression to the mean, a randomized controlled trial may be conducted to draw stronger conclusions in a future study. To gain additional insights into the effectiveness of the program, we can randomize patients into different treatment arms by providing different versions of the program. This could include varying the frequency or content of the lifestyle recommendations across the different treatment arms. Furthermore, we could investigate which lifestyle interventions, for example, increasing steps or improving sleep hygiene, result in greater BP improvements. With careful design, we can create a multiarm trial to investigate optimal engagement strategies and recommendations for different types of patients. Another limitation of this study is selection bias as the participants self-selected to enroll after receiving the recruitment flyer. To address this, we plan to recruit patients through PCP referrals. PCPs will refer their patients with high cardiovascular risk, who can benefit from our intervention. As previously mentioned, we hypothesize that this will increase the take-up rate due to increased trust from the more personal nature of the referral [39]. In addition, there is a need for a longer follow-up period as behavioral interventions can show improved outcomes during the first 6 months and then recidivism during the next 6 months. Finally, we did not collect socioeconomic data (eg, occupation, education, and income) from participants, preventing an analysis of how socioeconomic status impacts the program outcomes. In our future research, we will consider socioeconomic factors when analyzing the impact of the intervention. This analysis is imperative to ensure that the use of digital technologies does not contribute to an increased digital divide in health care and that all patients have equal access to high-quality health care [41,42].

Conclusions

To address the challenges of poor patient engagement due to generic, nonpersonalized lifestyle guidance and limited scalability of care due to human coaching models, we propose an AI-driven, autonomous, precise lifestyle coaching program for patients with hypertension. Patients who enrolled in the program experienced a significant improvement in BP. The program maintained a high engagement rate with minimal intervention from the care team. As the burden of hypertension increases globally, the necessity to develop new strategies to achieve hypertension control at scale is greater than ever. An AI-based, autonomous approach to hypertension-related lifestyle coaching can increase scalability and accessibility to effective BP management, ultimately improving the cardiovascular health of our community.

Acknowledgments

The research reported in this paper was supported by the Jacobs Family Endowed Chair fund for SD. The authors would like to thank Brian Khan for his clinical guidance during the study. The authors would like to thank Melissa Gellman, Russell Shimada, and Victoria Harris from the University of California, San Diego Health Population Health Services Organization for their support during the study.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to restrictions in the informed consent form.

Conflicts of Interest

JL, PHC, and SD are cofounders of CIPRA.ai Inc, a start-up company formed out of the University of California, San Diego, which has licensed the intervention technology presented in this paper. PA reports no conflict of interest.

Multimedia Appendix 1
Histogram showing the number of recommendations adhered to based on their difficulty rating. The average difficulty rating for recommendations that were followed was 1.97, indicating lower difficulty, whereas the average for those not followed was 3.67, indicating higher difficulty.
[PNG File , 109 KB - cardio_v8i1e51916_app1.png]

Multimedia Appendix 2
Participants' responses to the study experience survey. This survey asked patients to rate the difficulty level of completing the study tasks, how useful they found the recommendations, and their experience using the app.
[PNG File , 59 KB - cardio_v8i1e51916_app2.png]

References

1. Fryar CD, Ostchega Y, Hales CM, Zhang G, Kruszon-Moran D. Hypertension prevalence and control among adults: United States, 2015-2016. NCHS Data Brief 2017 Oct(289):1-8 [FREE Full text] [Medline: 29155682]
2. Hypertension cascade: hypertension prevalence, treatment and control estimates among US adults aged 18 years and older applying the criteria from the American College of Cardiology and American Heart Association's 2017 hypertension

guideline- NHANES 2017–2020. Centers for Disease Control and Prevention. 2017. URL: https://millionhearts.hhs.gov/data-reports/hypertension-prevalence.html [accessed 2023-05-12]
3. Lebeau J, Cadwallader J, Aubin-Auger I, Mercier A, Pasquet T, Rusch E, et al. The concept and definition of therapeutic inertia in hypertension in primary care: a qualitative systematic review. BMC Fam Pract 2014 Jul 02;15:130 [FREE Full text] [doi: 10.1186/1471-2296-15-130] [Medline: 24989986]
4. Kirkland EB, Heincelman M, Bishu KG, Schumann SO, Schreiner A, Axon RN, et al. Trends in healthcare expenditures among US adults with hypertension: national estimates, 2003-2014. J Am Heart Assoc 2018 May 30;7(11):e008731 [FREE Full text] [doi: 10.1161/JAHA.118.008731] [Medline: 29848493]
5. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison HC, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American College Of Cardiology/American Heart Association Task Force on clinical practice guidelines. Hypertension 2018 Jun;71(6):1269-1324. [doi: 10.1161/HYP.0000000000000066] [Medline: 29133354]
6. Blood Pressure Lowering Treatment Trialists' Collaboration. Pharmacological blood pressure lowering for primary and secondary prevention of cardiovascular disease across different levels of blood pressure: an individual participant-level data meta-analysis. Lancet 2021 May 01;397(10285):1625-1636 [FREE Full text] [doi: 10.1016/S0140-6736(21)00590-0] [Medline: 33933205]
7. Shimbo D, Artinian NT, Basile JN, Krakoff LR, Margolis KL, Rakotz MK, American Heart Association and the American Medical Association. Self-measured blood pressure monitoring at home: a joint policy statement from the American Heart Association and American Medical Association. Circulation 2020 Jul 28;142(4):e42-e63. [doi: 10.1161/CIR.0000000000000803] [Medline: 32567342]
8. Muntner P, Shimbo D, Carey RM, Charleston JB, Gaillard T, Misra S, et al. Measurement of blood pressure in humans: a scientific statement from the American Heart Association. Hypertension 2019 May;73(5):e35-e66 [FREE Full text] [doi: 10.1161/HYP.0000000000000087] [Medline: 30827125]
9. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: a systematic review and individual patient data meta-analysis. PLoS Med 2017 Sep;14(9):e1002389 [FREE Full text] [doi: 10.1371/journal.pmed.1002389] [Medline: 28926573]
10. Uhlig K, Patel K, Ip S, Kitsios GD, Balk EM. Self-measured blood pressure monitoring in the management of hypertension: a systematic review and meta-analysis. Ann Intern Med 2013 Aug 06;159(3):185-194 [FREE Full text] [doi: 10.7326/0003-4819-159-3-201308060-00008] [Medline: 23922064]
11. Appel LJ, Champagne CM, Harsha DW, Cooper LS, Obarzanek E, Elmer PJ, Writing Group of the PREMIER Collaborative Research Group. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. JAMA 2003;289(16):2083-2093. [doi: 10.1001/jama.289.16.2083] [Medline: 12709466]
12. Doughty KN, Del Pilar NX, Audette A, Katz DL. Lifestyle medicine and the management of cardiovascular disease. Curr Cardiol Rep 2017 Oct 04;19(11):116. [doi: 10.1007/s11886-017-0925-z] [Medline: 28980137]
13. Covassin N, Singh P. Sleep duration and cardiovascular disease risk: epidemiologic and experimental evidence. Sleep Med Clin 2016 Mar;11(1):81-89 [FREE Full text] [doi: 10.1016/j.jsmc.2015.10.007] [Medline: 26972035]
14. Cornelissen VA, Smart NA. Exercise training for blood pressure: a systematic review and meta-analysis. J Am Heart Assoc 2013 Feb 01;2(1):e004473 [FREE Full text] [doi: 10.1161/JAHA.112.004473] [Medline: 23525435]
15. Eckel RH, Jakicic JM, Ard JD, de Jesus JM, Houston Miller N, Hubbard VS, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 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 Jun 24;129(25 Suppl 2):S76-S99 [FREE Full text] [doi: 10.1161/01.cir.0000437740.48606.d1] [Medline: 24222015]
16. Bakris G, Ali W, Parati G. ACC/AHA versus ESC/ESH on hypertension guidelines: JACC guideline comparison. J Am Coll Cardiol 2019 Jun 18;73(23):3018-3026 [FREE Full text] [doi: 10.1016/j.jacc.2019.03.507] [Medline: 31196460]
17. Bell RA, Kravitz RL. Physician counseling for hypertension: what do doctors really do? Patient Educ Couns 2008 Jul;72(1):115-121. [doi: 10.1016/j.pec.2008.01.021] [Medline: 18328663]
18. Wilson-Anumudu F, Quan R, Cerrada C, Juusola J, Castro Sweet C, Bradner Jasik C, et al. Pilot results of a digital hypertension self-management program among adults with excess body weight: single-arm nonrandomized trial. JMIR Form Res 2022 Mar 30;6(3):e33057 [FREE Full text] [doi: 10.2196/33057] [Medline: 35353040]
19. Toro-Ramos T, Kim Y, Wood M, Rajda J, Niejadlik K, Honcz J, et al. Efficacy of a mobile hypertension prevention delivery platform with human coaching. J Hum Hypertens 2017 Dec;31(12):795-800. [doi: 10.1038/jhh.2017.69] [Medline: 28972573]
20. Milani RV, Lavie CJ, Bober RM, Milani AR, Ventura HO. Improving hypertension control and patient engagement using digital tools. Am J Med 2017 Jan;130(1):14-20. [doi: 10.1016/j.amjmed.2016.07.029] [Medline: 27591179]
21. Mao AY, Chen C, Magana C, Caballero Barajas K, Olayiwola JN. A mobile phone-based health coaching intervention for weight loss and blood pressure reduction in a national payer population: a retrospective study. JMIR Mhealth Uhealth 2017 Jun 08;5(6):e80 [FREE Full text] [doi: 10.2196/mhealth.7591] [Medline: 28596147]

22. Branch OH, Rikhy M, Auster-Gussman LA, Lockwood KG, Graham SA. Relationships between blood pressure reduction, weight loss, and engagement in a digital app-based hypertension care program: observational study. JMIR Form Res 2022 Oct 27;6(10):e38215 [FREE Full text] [doi: 10.2196/38215] [Medline: 36301618]
23. Leitner J, Chiang P, Khan B, Dey S. An mHealth lifestyle intervention service for improving blood pressure using machine learning and IoMTs. In: Proceedings of the 2022 IEEE International Conference on Digital Health. 2022 Presented at: ICDH '22; July 10-16, 2022; Barcelona, Spain p. 142-150 URL: https://ieeexplore.ieee.org/document/9861082 [doi: 10.1109/icdh55609.2022.00030]
24. Monitoring your blood pressure at home. American Heart Association. URL: https://www.heart.org/en/health-topics/high-blood-pressure/understanding-blood-pressure-readings/monitoring-your-blood-pressure-at-home [accessed 2024-04-05]
25. Siervo M, Lara J, Chowdhury S, Ashor A, Oggioni C, Mathers JC. Effects of the Dietary Approach to Stop Hypertension (DASH) diet on cardiovascular risk factors: a systematic review and meta-analysis. Br J Nutr 2015 Jan 14;113(1):1-15. [doi: 10.1017/S0007114514003341] [Medline: 25430608]
26. Henriksen A, Haugen Mikalsen M, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables. J Med Internet Res 2018 Mar 22;20(3):e110 [FREE Full text] [doi: 10.2196/jmir.9157] [Medline: 29567635]
27. Germini F, Noronha N, Borg Debono V, Abraham Philip B, Pete D, Navarro T, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: systematic review of the literature. J Med Internet Res 2022 Jan 21;24(1):e30791 [FREE Full text] [doi: 10.2196/30791] [Medline: 35060915]
28. Miller DJ, Sargent C, Roach GD. A validation of six wearable devices for estimating sleep, heart rate and heart rate variability in healthy adults. Sensors (Basel) 2022 Aug 22;22(16):6317 [FREE Full text] [doi: 10.3390/s22166317] [Medline: 36016077]
29. Yuenyongchaiwat K, Pipatsitipong D, Sangprasert P. Increasing walking steps daily can reduce blood pressure and diabetes in overweight participants. Diabetol Int 2018 Feb;9(1):75-79 [FREE Full text] [doi: 10.1007/s13340-017-0333-z] [Medline: 30603352]
30. Lefferts EC, Saavedra JM, Song BK, Brellenthin AG, Pescatello LS, Lee D. Increasing lifestyle walking by 3000 steps per day reduces blood pressure in sedentary older adults with hypertension: results from an e-health pilot study. J Cardiovasc Dev Dis 2023 Jul 27;10(8):317 [FREE Full text] [doi: 10.3390/jcdd10080317] [Medline: 37623330]
31. Bock JM, Vungarala S, Covassin N, Somers VK. Sleep duration and hypertension: epidemiological evidence and underlying mechanisms. Am J Hypertens 2022 Jan 05;35(1):3-11 [FREE Full text] [doi: 10.1093/ajh/hpab146] [Medline: 34536276]
32. Ali W, Gao G, Bakris GL. Improved sleep quality improves blood pressure control among patients with chronic kidney disease: a pilot study. Am J Nephrol 2020;51(3):249-254. [doi: 10.1159/000505895] [Medline: 31982868]
33. Ponte Márquez PH, Feliu-Soler A, Solé-Villa MJ, Matas-Pericas L, Filella-Agullo D, Ruiz-Herrerias M, et al. Benefits of mindfulness meditation in reducing blood pressure and stress in patients with arterial hypertension. J Hum Hypertens 2019 Mar;33(3):237-247. [doi: 10.1038/s41371-018-0130-6] [Medline: 30425326]
34. Gupta DK, Lewis CE, Varady KA, Su YR, Madhur MS, Lackland DT, et al. Effect of dietary sodium on blood pressure: a crossover trial. JAMA 2023 Dec 19;330(23):2258-2266. [doi: 10.1001/jama.2023.23651] [Medline: 37950918]
35. He FJ, Tan M, Ma Y, MacGregor GA. Salt reduction to prevent hypertension and cardiovascular disease: JACC state-of-the-art review. J Am Coll Cardiol 2020 Feb 18;75(6):632-647 [FREE Full text] [doi: 10.1016/j.jacc.2019.11.055] [Medline: 32057379]
36. Alagappan R, Das S. Uncovering Twilio: insights into cloud communication services. University of Wisconsin System. URL: https://pages.cs.wisc.edu/~ra/Twilio.pdf [accessed 2024-04-05]
37. Fogg BJ. Fogg behavior model. Behavior Design Lab. URL: https://behaviormodel.org [accessed 2024-04-05]
38. Franklin SS, Thijs L, Hansen TW, O'Brien E, Staessen JA. White-coat hypertension: new insights from recent studies. Hypertension 2013 Dec;62(6):982-987. [doi: 10.1161/HYPERTENSIONAHA.113.01275] [Medline: 24041952]
39. Pfaff E, Lee A, Bradford R, Pae J, Potter C, Blue P, et al. Recruiting for a pragmatic trial using the electronic health record and patient portal: successes and lessons learned. J Am Med Inform Assoc 2019 Jan 01;26(1):44-49 [FREE Full text] [doi: 10.1093/jamia/ocy138] [Medline: 30445631]
40. Bundy JD, Li C, Stuchlik P, Bu X, Kelly TN, Mills KT, et al. Systolic blood pressure reduction and risk of cardiovascular disease and mortality: a systematic review and network meta-analysis. JAMA Cardiol 2017 Jul 01;2(7):775-781 [FREE Full text] [doi: 10.1001/jamacardio.2017.1421] [Medline: 28564682]
41. Ramsetty A, Adams C. Impact of the digital divide in the age of COVID-19. J Am Med Inform Assoc 2020 Jul 01;27(7):1147-1148 [FREE Full text] [doi: 10.1093/jamia/ocaa078] [Medline: 32343813]
42. Rodriguez JA, Clark CR, Bates DW. Digital health equity as a necessity in the 21st century cures act era. JAMA 2020 Jun 16;323(23):2381-2382. [doi: 10.1001/jama.2020.7858] [Medline: 32463421]

Abbreviations

AI: artificial intelligence
API: Application Programming Interface
BP: blood pressure

DBP: diastolic blood pressure
ML: machine learning
PCP: primary care physician
PHSO: Population Health Services Organization
SBP: systolic blood pressure

Edited by A Mavragani; submitted 24.08.23; peer-reviewed by Y Jeem, J Cheng, P Washington; comments to author 14.11.23; revised version received 29.01.24; accepted 09.04.24; published 28.05.24.

Please cite as:
Leitner J, Chiang PH, Agnihotri P, Dey S
The Effect of an AI-Based, Autonomous, Digital Health Intervention Using Precise Lifestyle Guidance on Blood Pressure in Adults With Hypertension: Single-Arm Nonrandomized Trial
JMIR Cardio 2024;8:e51916
URL: https://cardio.jmir.org/2024/1/e51916
doi:10.2196/51916
PMID:38805253

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