JMIR Cardio

Contents

Original Papers

Original Paper

Accuracy and Usability of a Novel Algorithm for Detection of Irregular Pulse Using a Smartwatch Among Older Adults: Observational Study

Eric Y Ding[1], MSc; Dong Han[2], BSc; Cody Whitcomb[3], BSc; Syed Khairul Bashar[2], BSc; Oluwaseun Adaramola[3], MD; Apurv Soni[1], BSc; Jane Saczynski[4], PhD; Timothy P Fitzgibbons[3], MD, PhD; Majaz Moonis[5], MD; Steven A Lubitz[6], MD, MPH; Darleen Lessard[1], MSc; Mellanie True Hills[7], BSc; Bruce Barton[1], PhD; Ki Chon[2], PhD; David D McManus[3], MD, MSci

[1]Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

[2]Department of Biomedical Engineering, University of Connecticut, Storrs, CT, United States

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

[4]Department of Pharmacy and Health Systems Sciences, Northeastern University, Boston, MA, United States

[5]Department of Neurology, University of Massachusetts Medical School, Worcester, MA, United States

[6]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, United States

[7]StopAfib.org, American Foundation for Women's Health, Decatur, TX, United States

Corresponding Author:
Eric Y Ding, MSc
Department of Population and Quantitative Health Sciences
University of Massachusetts Medical School
368 Plantation Street
Albert Sherman Center, AS7-1046
Worcester, MA, 01605
United States
Phone: 1 4086607644
Email: eric.ding@umassmed.edu

Abstract

Background: Atrial fibrillation (AF) is often paroxysmal and minimally symptomatic, hindering its diagnosis. Smartwatches may enhance AF care by facilitating long-term, noninvasive monitoring.

Objective: This study aimed to examine the accuracy and usability of arrhythmia discrimination using a smartwatch.

Methods: A total of 40 adults presenting to a cardiology clinic wore a smartwatch and Holter monitor and performed scripted movements to simulate activities of daily living (ADLs). Participants' clinical and sociodemographic characteristics were abstracted from medical records. Participants completed a questionnaire assessing different domains of the device's usability. Pulse recordings were analyzed blindly using a real-time realizable algorithm and compared with gold-standard Holter monitoring.

Results: The average age of participants was 71 (SD 8) years; most participants had AF risk factors and 23% (9/39) were in AF. About half of the participants owned smartphones, but none owned smartwatches. Participants wore the smartwatch for 42 (SD 14) min while generating motion noise to simulate ADLs. The algorithm determined 53 of the 314 30-second noise-free pulse segments as consistent with AF. Compared with the gold standard, the algorithm demonstrated excellent sensitivity (98.2%), specificity (98.1%), and accuracy (98.1%) for identifying irregular pulse. Two-thirds of participants considered the smartwatch highly usable. Younger age and prior cardioversion were associated with greater overall comfort and comfort with data privacy with using a smartwatch for rhythm monitoring, respectively.

Conclusions: A real-time realizable algorithm analyzing smartwatch pulse recordings demonstrated high accuracy for identifying pulse irregularities among older participants. Despite advanced age, lack of smartwatch familiarity, and high burden of comorbidities, participants found the smartwatch to be highly acceptable.

XSL•FO
RenderX

Introduction

Background

Atrial fibrillation (AF) is the most common heart rhythm problem in the world and the number of patients living with AF is increasing rapidly [1,2]. The arrhythmia confers a 3-fold higher risk for heart failure, a 2-fold risk for dementia, and a 5-fold risk of ischemic stroke among affected individuals, irrespective of the symptom severity or pattern [1,3]. The diagnosis of AF often represents a clinical challenge, especially in its early stages, owing to its paroxysmal and sometimes asymptomatic nature. Despite the fact that short-lived episodes may elude clinical surveillance, brief or infrequent episodes of AF remain associated with higher risk for stroke and death [4,5]. Treatment of AF with oral anticoagulation drastically reduces the risk for stroke by up to 70% [6], but many patients escape detection until after a serious complication. For example, 1 in 5 AF patients present with stroke as their first manifestation of the arrhythmia [7]. New methods for monitoring and screening are needed to facilitate early AF diagnosis, initiate treatment, and reduce suffering and death from the arrhythmia.

Systematic and opportunistic screening of older populations for AF using mobile and digital health is feasible and can identify asymptomatic, community-dwelling individuals with undiagnosed AF [8]. Recent European Society of Cardiology AF guidelines emphasize the importance of opportunistic screening but stop short of recommending systematic screening for all individuals, in part owing to the expensive and sometimes cumbersome nature of the existing rhythm monitoring strategies [9]. There is considerable interest, however, from AF patients and families, AF patient advocacy groups, health care providers, insurers, and health systems to develop lower-cost, more user-friendly solutions for long-term heart rhythm monitoring using mobile health (mHealth) devices to empower patients and reduce suffering from AF [10-12]. The Apple Watch Series 4 recently received Food and Drug Administration (FDA) clearance for its mobile electrocardiogram (ECG) function that has been validated in AF patients (data yet unpublished at the time of writing) [13], marking a monumental milestone in smartwatch AF monitoring and other mobile technology developers are likely to follow suit in the near future. Professional organizations such as the American Heart Association (AHA) have also shown great enthusiasm and support for these ventures, and the current AHA president Dr Ivor Benjamin has claimed that smartwatches that "capture data [...] in real time [are] changing the way we practice medicine" [14]. We and other groups have developed accurate automated, real-time realizable, pulse-based approaches for AF detection using a smartphone, but this approach, as with ECG-based approaches using devices that pair with smartphones (ie, AliveCor Kardia [15]), require that a participant perform an active rhythm check [10,11]. As prior studies have demonstrated that even minutes of AF confer risk for ischemic stroke [16], and as smartwatches can perform passive, frequent pulse assessments over long periods of time [17,18], there is considerable interest in adapting existing pulse-based approaches for use on a smartwatch [17,19,20].

Although appealing, the use of a smartwatch for AF monitoring introduces unique and significant technical challenges, such as motion and noise artifacts generated during activities of daily living (ADLs), as well as usability concerns, as individuals at risk for AF tend to be older, less familiar with mHealth devices, and frequently affected by physical and cognitive impairments that can impede operation of, and comfort with, mHealth technologies [21].

Objectives

In this investigation, we sought to test the performance of a novel, real-time realizable, automated algorithm for AF discrimination using pulse data obtained from a smartwatch among older individuals with, or at risk for, AF while they executed simulated ADLs. Furthermore, we assessed study participants' impressions of the smartwatch, generally and across specific usability domains. Finally, we identified characteristics associated with comfort using a smartwatch for rhythm analysis.

Methods

Design and Setting

This observational study was designed to evaluate the performance and usability of a smartwatch for heart rhythm analysis among older individuals with, or at high risk for, AF. Participants were enrolled between June 2016 and November 2017 from the ambulatory clinics at the University of Massachusetts Medical Center. All participants provided written informed consent before the study participation. This study was approved by the University of Massachusetts Medical School (UMMS) Institutional Review Board (UMMS IRB number H00009953).

Study Population

Study staff reviewed the electronic health records (EHR) of all patients presenting for an ambulatory visit. To be considered eligible for enrollment, participants were required to be 21 years of age or older, capable of and willing to provide informed consent, and able to speak and read English. Individuals were excluded from study participation if they were pregnant, incarcerated, had reported an adverse reaction to ECG electrodes or a Holter monitor, or refused to adhere to any aspect of the proposed study protocol, including a brief walk test. The staff telephoned potentially eligible study participants (both those with and without AF) 1 to 2 weeks before their clinic visit to assess interest in study participation. A total of 78 patients were telephoned and 48 patients expressed interest in the study. These patients were then approached for consent after their clinic visit. However, 7 patients declined participation at this time and 1 patient was excluded due to wrist size being too large for the smartwatch, resulting in a final sample of 40 participants.

Trained staff abstracted clinical, electrocardiographic, and laboratory data from the EHR on all participants, including data

XSL•FO

RenderX

obtained during the ambulatory visit immediately preceding the study examination. Resting heart rate and rhythm status, as well as vital signs, including respiratory rate, systolic and diastolic blood pressure, and body mass index, were obtained on all participants. The use of cardiovascular medications was also abstracted from the EHR. The wrist circumference and skin tone of the participants were determined by the research staff at the time of their study examination as these factors may potentially influence pulse recordings obtained by the smartwatch.

Study Procedures to Simulate Activities of Daily Living

After providing a brief overview of the smartwatch (Samsung Simband 2, Figure 1) and the study protocol, the trained staff helped the participants to don the smartwatch and wear the 7-lead Holter monitor (Rozinn RZ153+ Series, Rozinn Electronics Inc). After ensuring that the watch and Holter monitor were properly fitted and actively recording, participants were asked to sit still for 2 min, followed by a 2-min slow walk (2 miles per hour) down the clinic hallway, followed by a 30-second standing period. The study participants were then instructed to walk quickly (4 miles per hour) down the clinic hallway for 2 min, followed by, in sequence, a 1-min stand, a 1-min period of vertical watch-arm movements, a 1-min period of watch-arm wrist movements, and a 30-second period of standing still [22]. Participants were then asked to sit and stand repeatedly from a chair over a 1-min period. Participants were then asked to climb and descend stairs over a 2-min period, followed by a 1-min period where participants sat and performed

deep breathing exercises [23,24]. Slight modifications to the protocol were made for 3 participants to ensure safety: the first participant omitted the repeated sit and stand sequence because of recent knee injury, the second only completed 1 min of climbing stairs because of exertional dyspnea, and the third participant declined the stairs climbing portion of the protocol because of injury from a recent motor vehicle accident. Participants also wore the watch while completing the questionnaire.

Signal Acquisition, Transfer, and Blinding Procedures

The Samsung Simband 2 (Figure 1) is a wrist-worn mHealth device capable of performing continuous real-time monitoring of biophysical data, providing real-time user feedback and wireless, secure, and asynchronous signal transfer [25]. When connected to a secure wireless internet network, the Samsung Simband 2 passively uploads recorded data to ARTIK Cloud (Samsung Electronics), a Web-based cloud-based data storage platform used for research. For this study, Samsung engineers provided us with access to ARTIK Cloud and technical support. We generated a study identification number (ID) for each participant and entered this into the Simband 2 to link photoplethysmogram (PPG) pulse data to secure study data (Figure 2). PPG data identifiable only based on study ID was uploaded to the ARTIK Cloud using a secure connection. Investigators at the University of Connecticut (UConn) group who were blinded to the participants' rhythm status performed offline rhythm analysis using the real-time realizable algorithm.

Figure 1. Samsung Simband 2 smartwatch showing simultaneous single-lead ECG (electrocardiogram) and PPG (photoplethysmogram) recordings.

Figure 2. Representative electrocardiographic, pulse, and pulse interval waveforms as recorded by the Samsung Simband 2.0 smartwatch from study participants in various rhythm states. Panels 1 to 4 (top to bottom) represent patients in various rhythm states: (1) shows a patient in normal sinus rhythm, (2) shows a patient with premature atrial contractions, (3) shows a patient with premature ventricular contractions, and (4) shows a patient in atrial fibrillation. For each patient, the single-lead electrocardiogram (a) and pulse plethysmography waveforms (b) are collected by the smartwatch. The pulse interval (c) and Poincare plots (graphed on right) for each patient are also calculated and represented. BPM: beats per min; HR: heart rate; ECG: electrocardiogram; PPG: photoplethysmogram.
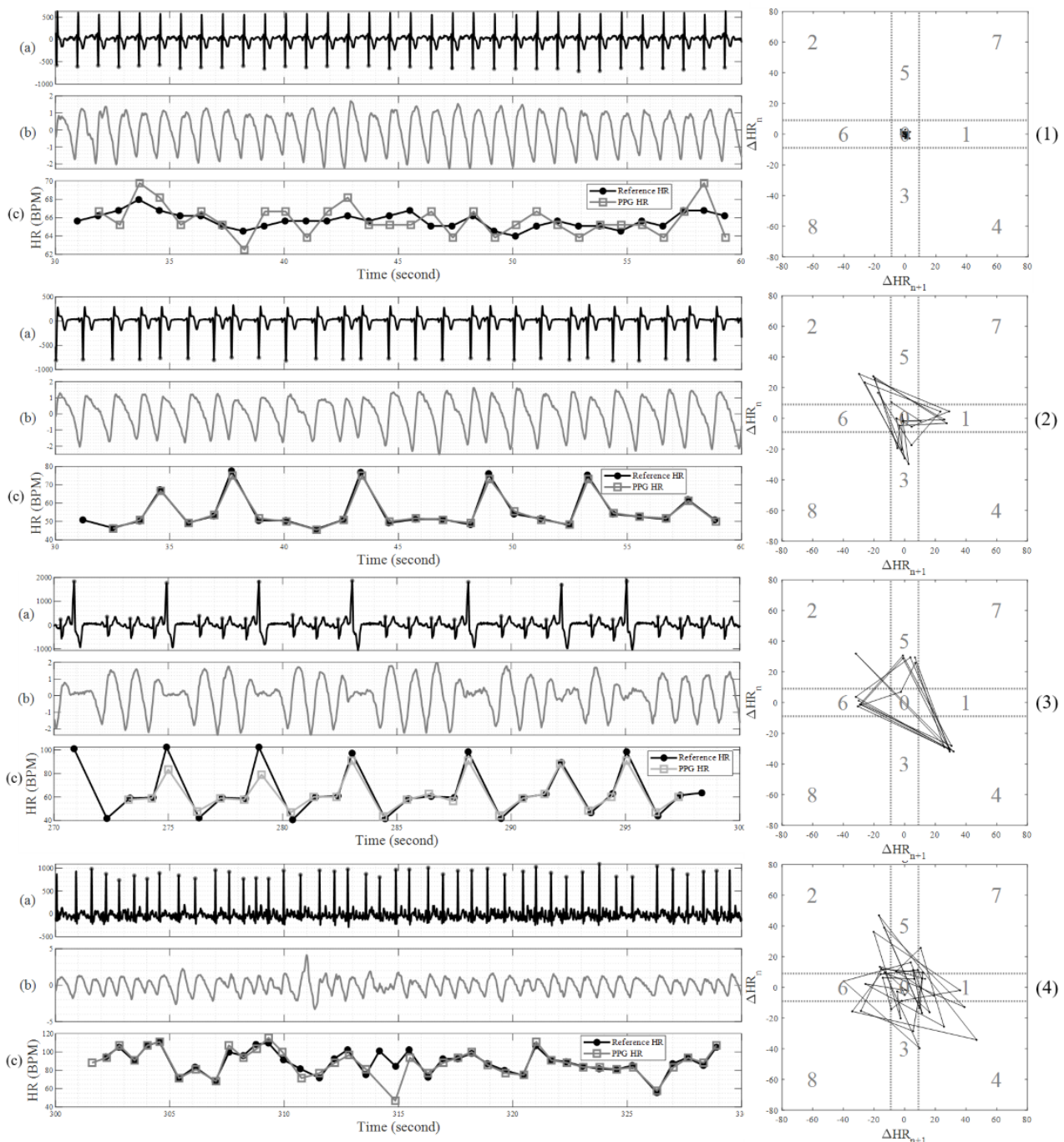
Motion Noise Detection

Once downloaded onto secure UConn study servers, each participant's pulse data were divided into 30-second segments and then analyzed using a novel motion noise artifact (MNA) detection algorithm. The details of the MNA detection algorithm have been described elsewhere [26]. In brief, the MNA algorithm employs a combination of features derived from a time-frequency analysis and the significance of the accelerometer data's amplitude to determine MNA severity and a previously derived threshold MNA value is used to determine whether or not a segment is analyzable [26]. After discarding MNA-corrupted PPG segments, we calculated features to classify AF from normal sinus rhythm (NSR).

Pulse Analysis and Rhythm Discrimination

AF is characterized by disorganized atrial electrical activity that stimulates the ventricles in a random fashion that increases beat-to-beat variability. Our approach to pulse analysis from PPG data has been described in detail and uses 2 validated

statistical techniques. The first of these methods, sample entropy (SampEn), measures the complexity of pulse variability. The second, root mean square of successive difference of RR intervals (RMSSD), quantifies peak-to-peak pulse waveform variability; the pulse peak determination methodology has also been previously described [27]. Next, SampEn and RMSSD are combined into a single parameter (Comb) to discriminate between AF and normal rhythms. The combined parameter yields greater area under the receiver operating characteristic curve than either SampEn or RMSSD. We used the combined threshold value of Comb=0.94. Moreover, as a second step, a novel method based on Poincare plot was used to identify ectopic premature atrial or ventricular contractions (PACs, PVCs; paper under review and personal communication by DH, September 2018). Each 30-second segment of PPG data obtained from the smartwatch was blindly assigned, by the UConn team, a designation of *AF* or *not AF* on the basis of this automated analysis. We have previously derived and validated thresholds for detecting PACs/PVCs using a 2-dimensional Poincare plot, which plots the length of each RR interval against the previous RR interval to visualize the beat-to-beat variability in these waveforms among individuals with an irregular pulse [28]. This graph is used to identify bigeminal, trigeminal, and quadrigeminal patterns, which are frequently observed patterns among patients with PACs and PVCs and are common sources of false positive AF detection [28].

Gold-Standard Electrocardiogram-Based Rhythm Analysis

A 7-lead Holter monitor was used to record an ECG as the gold standard for comparison. ECG data from each participant were labeled with the participant's study ID and recorded on an Secure Digital card. Staff downloaded the ECG data and transferred it securely to UConn investigators (KC, DH, SKB), where it was stored on firewall protected servers and analyzed separately from the pulse data from watches. Holter data were obtained for 41 participants, but were unavailable for 1 participant, likely because of inadequate contact of leads.

The ECG data were divided into 30-second segments and linked to PPG data using time-stamps and study IDs. Each segment of ECG data was analyzed by a highly accurate and well validated AF detection algorithm using a combination of time-varying coherence functions and Shannon Entropy (ShE) [29] to establish the rhythm status (gold-standard) for each 30-second time period. This algorithm is currently used by ScottCare CardioView Dx, a diagnostic partner program for Holter and event monitors designed by ScottCare Cardiovascular Solutions. Similar to the pulse waveforms obtained from the smartwatch, a subsequent ectopic beat identification algorithm using a Poincare plot threshold was also applied to the ECG data windows deemed to have irregular pulse. For quality control, a board-certified cardiac electrophysiologist (DDM) blinded to the results of the automated ECG analysis reviewed 10% of the 30-second ECG recordings deemed to be consistent with AF and 10% of those deemed to be sinus rhythm. Consistent with prior reports demonstrating excellent performance for AF detection, there was 100% agreement between clinician and automated ECG rhythm determinations [29].

Study Questionnaire

We characterized system usability, as well as the psychosocial, cognitive, and sociodemographic characteristics of study participants using validated instruments. We also assessed other factors known to influence perceptions of mHealth devices, including education level, yearly income, employment status, and prior experience with smart devices, as measured by smartphone or smartwatch ownership and social media use.

Usability

Usability in this study was measured globally and across several usability domains. The Brooke System Usability Scale (SUS) is a widely used and validated 10-item questionnaire that assesses multiple dimensions of usability and has been used to assess prior mHealth devices [30]. Responses to each question were scored using a standardized 5-point Likert scale (1=strongly disagree; 5=strongly agree). Responses were weighted equally and converted to achieve a maximum score of 100. For reference, a score of ≥68 is consistent with high usability [30].

A separate investigator-generated assessment was also administered to measure participants' perceptions of the smartwatch's ease of use, its overall importance, privacy concerns related to smartwatch use, perceived fit of the device into daily activities, and comfort with use, as measured by stress associated with using the device. Participants responded to all questions using the same 5-point Likert scale used in Brooke SUS.

Cognitive Impairment and Mood

We employed 3 validated questionnaires to assess cognitive impairment and mood. The Montreal Cognitive Assessment is a validated tool for screening of mild cognitive impairment and is widely used in clinical settings [31]. It includes 30 items and assesses multiple cognitive domains. The scoring ranges from 0 to 30 and a score <26 is indicative of mild cognitive impairment. The Patient Health Questionnaire–9 is a validated 9-item instrument for screening, diagnosing, and monitoring depression in the clinic [32]. A score of 5 to 9 is indicative of mild, 10 to 14 of moderate, and greater than 15 of severe depression. The Generalized Anxiety Disorder scale is a multipurpose 7-item instrument to describe the severity of the patient's anxiety symptoms in the past 2 weeks [33]. It is a well-validated tool used to screen for a variety of anxiety related disorders, such as generalized anxiety disorder, panic disorder, and posttraumatic stress disorder. The scores range from 0 to 21, and a score of 5 to 9 is considered mild anxiety, 10 to 14 as moderate anxiety or related condition, and greater than 15 as severe anxiety or related condition. This section of the questionnaire was completed by 22 out of the 40 study participants.

Data Analysis

We first calculated the proportion of noise-free data segments out of the total number of data segments. We then calculated the appropriate test characteristics, including sensitivity, specificity, and accuracy, to examine the performance of our automated pulse analysis algorithm from 314 noise-free pulse segments for the detection of an irregular pulse consistent with

AF compared with the results of a validated algorithm examining contemporaneous 7-lead Holter ECG (criterion standard), using established threshold values of RMSSD, SampEn, and Poincare plot [10,11].

The overall usability of the smartwatch for arrhythmia monitoring was examined using validated Brooke SUS. Unadjusted linear regression was then used to identify patient-level characteristics associated with overall Brooke SUS score, as well as Likert-type scores across several usability domains (system ease of use, system importance to user, system privacy concerns, perceived fit of system into daily activities, and comfort with the system, as measured by stress induced by use). Age, sex, history of coronary artery disease, coronary bypass graft procedure, cardioversion, stroke, education level, smartphone ownership, social media use, cognitive impairment, depression, and anxiety were examined as predictors of these usability outcomes. All analyses were performed in MATLAB 9.1 (MathWorks) and Stata 13 (StataCorp).

Results

Participant Characteristics

The characteristics of the 40 participants enrolled in the study are shown in Table 1. The average age of the participants was 71 (SD 8) years, 80% (32/40) were male, and all were Caucasian. About 1 in 3 participants had a history of coronary artery disease, 35% (14/40) had undergone a previous cardioversion, and 23% (9/39) were in AF at the time of their study visit. Cognitive impairment was common (59%, 13/22), but symptoms of mild anxiety (18%, 4/22) and depression (18%, 4/22) were less frequently observed. Most participants were retired (70%, 28/40) and most participants reported an annual income of US $50,000, Can $67,404.25, Aus $71,560 or more. Most participants had achieved a high level of education, with over half having completed college or a graduate degree program. Although more than half of the participants owned a smartphone (55%, 22/40), none owned a smartwatch at the time of the study visit. Participants in AF at the time of their study visit had, on average, higher heart rates, as well as higher rates of prior heart failure, cardioversion, anticoagulant use, and anxiety.

Pulse Rhythm Analysis

A total of 40 participants wore the smartwatch for an average of 42 (SD 14) min, generating a total of 2538 30-second data segments of pulse waveform recordings, of which 314 were noise-free. All data from 1 participant were corrupted by motion/noise artifact, likely from a poor wristband fit. Furthermore, 63 out of the 314 clean 30-second pulse segments were deemed to be *irregular/consistent with AF* and 251 were determined to be *regular/consistent with normal rhythm*. All windows designated irregular pulse were subsequently subjected to ectopic beat detection, which correctly salvaged 6 windows of benign beat irregularities as such compared with ECG data. Final analysis of the smartwatch-generated pulse segments resulted in 54 windows deemed AF, 6 windows deemed PACs/PVCs, and 248 windows deemed NSR. Compared with the ECG rhythm gold standard, AF was correctly identified from smartwatch pulse data in 54 out of 55 windows, corresponding to a sensitivity of 98.2%. The algorithm correctly identified 254 out of 259 windows as showing either normal rhythm or PACs/PVCs compared with the reference standard, corresponding to 98.1% specificity. The pulse analysis algorithm demonstrated excellent overall accuracy (98.1%) in detecting the presence of AF. The algorithm had a positive predictive value of 91.5% and a negative predictive value of 99.6%.

Usability Analysis

All 40 participants were included in the usability analysis. The smartwatch demonstrated high usability for rhythm analysis, as determined using the validated Brooke SUS, with over two-thirds of participants (67.7%) considering the watch to be highly usable. The average Brooke SUS score was 72.9 (SD 17.5). Individual usability domains, including ease of use, importance to user, fit into daily activity, comfort with privacy, and stress associated with use were also assessed using a Likert scale (1 to 5). These usability domains were generated by the investigators for the specific purpose of evaluating this technology and are presented as univariate dot plots to best represent distribution of responses, which may be more meaningful than summary statistics. Overall usability was high across usability domains (Figure 3), with no significant variation across areas. Unadjusted regression analysis (Table 2) showed that older age and prior history of coronary artery bypass surgery were associated with less comfort (as measured by greater stress), when using the smartwatch for rhythm analysis (unadjusted beta coefficients –0.039 and –0.935, respectively). The history of having undergone a cardioversion for AF was associated with greater comfort with sharing rhythm data using the smartwatch (unadjusted beta coefficient 0.72).

Table 1. Baseline characteristics of study participants (N=40).

Characteristics	Statistics
Demographic characteristics	
Age (years), mean (SD)	70.6 (8)
Sex, male, n (%)	32 (80)
Race, Caucasian, n (%)	40 (100)
Skin tone: tan	27 (68)
Skin tone: pale	10 (25)
Skin tone: unspecified	3 (8)
Wrist circumference, inches, mean (SD), (n=32)	6.9 (0.7)
Medical characteristics	
$CHA_2 DS_2$-VASc score[a], mean (SD)	2.6 (1.3)
Body mass index, kg/m^2, mean (SD)	29.3 (5.1)
Respiratory rate, bpm[b], mean (SD)	16.7 (1.2)
Resting heart rate (per electrocardiogram), bpm, mean (SD)	68.8 (14.3)
Systolic blood pressure, mm Hg, mean (SD)	126.3 (17.7)
Diastolic blood pressure, mm Hg, mean (SD)	72.0 (10.4)
Hypertension, n (%)	25 (63)
Hyperlipidemia, n (%)	23 (58)
Current smoking, n (%)	0 (0)
Diabetes mellitus, Type 2, n (%)	9 (23)
Coronary artery disease, n (%)	13 (33)
Prior coronary artery bypass graft, n (%)	7 (18)
Congestive heart failure, n (%)	2 (5)
Sleep apnea, n (%)	10 (25)
Prior cardioversion (%)	14 (35)
Stroke, n (%)	2 (5)
Arrhythmia characteristics, n (%)	
History of atrial fibrillation	28 (70)
Type of atrial fibrillation	
Paroxysmal	17 (60)
Permanent	4 (14)
Persistent	5 (18)
Unspecified	2 (7)
Rhythm at time of pulse assessment	
Sinus rhythm	30 (77)
Atrial fibrillation	9 (23)
Treatment characteristics, n (%)	
Beta-blocker	27 (68)
Calcium channel blocker	12 (30)
Statin	29 (73)
Antiarrhythmic drug	12 (30)
Digoxin	1 (3)

Characteristics	Statistics
Anticoagulant	21 (53)
Psychosocial characteristics[c], n (%)	
Cognitive impairment	13 (59)
Anxiety	4 (18)
Depression	4 (18)
Educational status, n (%)	
Completed some high school	2 (5)
Graduated high school	9 (23)
Graduated high school, some college	5 (13)
Graduated college	8 (20)
Graduated college, some graduate school	2 (5)
Completed a graduate degree	14 (35)
Employment status, n (%)	
Employed full-time	7 (17)
Employed part-time	4 (10)
Full time home-maker or caretaker	1 (3)
Retired	28 (70)
Income status, n (%)	
Less than $10,000	1 (4)
$10,000 to $29,999	1 (4)
$30,000 to $49,999	3 (11)
$50,000 to $69,000	3 (11)
$70,000 to $89,999	8 (30)
$90,000 to $149,999	9 (33)
$150,000 or more	2 (7)
Unreported	13 (33)
Technology use, n (%)	
Own smartphone	22 (55)
Own smart watch	0 (0)

[a]CHA_2DS_2-VASc score: clinically used tool for stroke risk assessment.

[b]bpm: beats per minute.

[c]Cognitive impairment is based on Montreal Cognitive Assessment score <26, anxiety is based on Generalized Anxiety Disorder–7 score >4, and depression is based on Patient Health Questionnaire–9 score >4. Data are available for 22 participants for all measures.

Figure 3. Responses to smartwatch for atrial fibrillation usability questions. Each circle represents an individual participant's coded response.
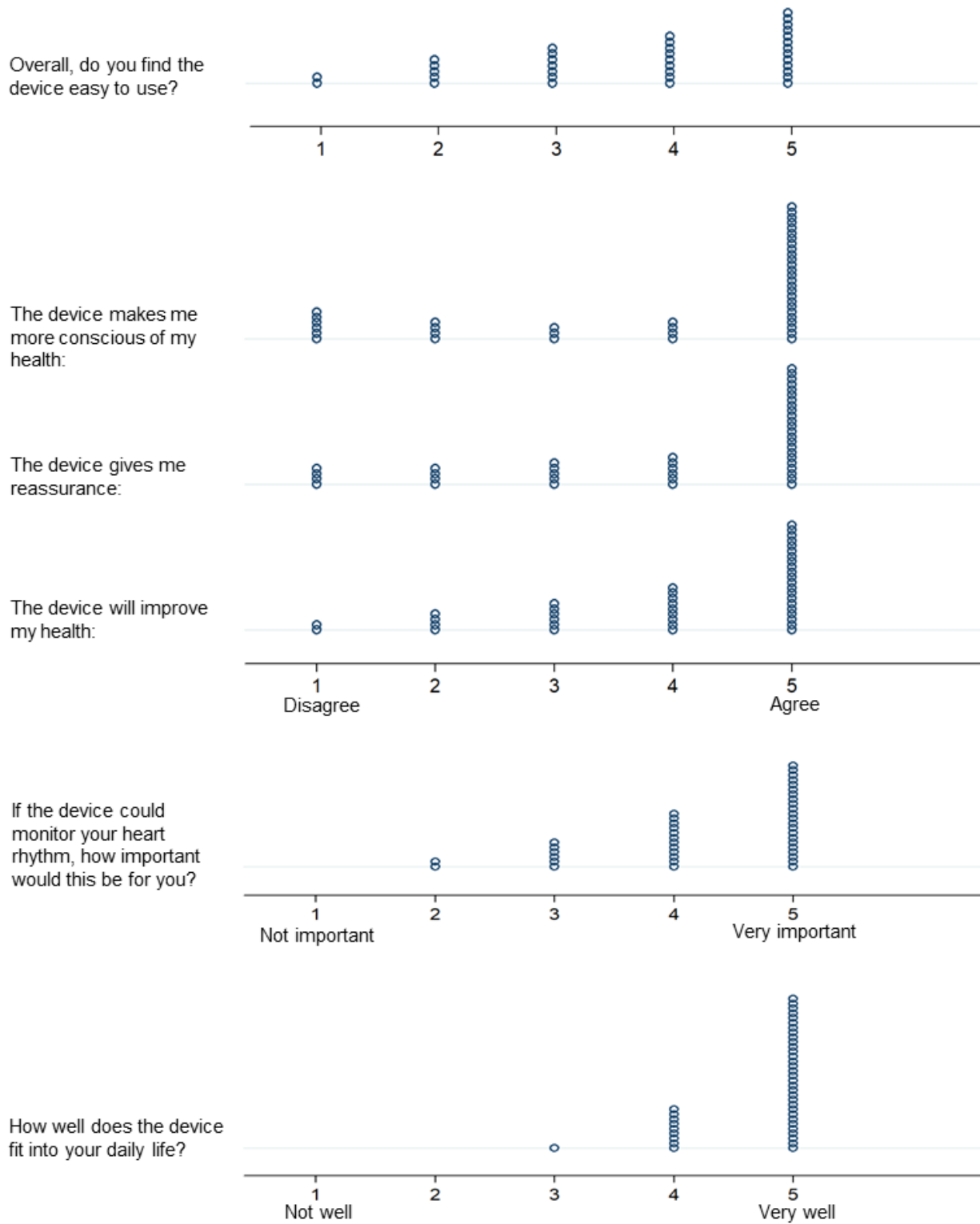
Table 2. Factors associated with individual usability domains and system usability score.

Variables	Individual usability domains					Overall Brooke system usability score[b]
	Ease of use[b]	Importance to user[b]	Fit into daily activity[b]	Comfort with privacy[b]	Comfort with use[b]	
Age	0.02	–0.01	–0.02	0.00	–0.039[a]	–0.082
Sex, male	–0.38	–0.01	0.14	0.52	–0.063	–8.905
History of coronary artery disease[c]	0.86	0.13	–0.19	–0.23	0.046	–2.680
Prior coronary artery bypass graft[c]	–0.54	–0.04	–0.26	0.21	–0.935[a]	–2.857
Prior cardioversion[c]	–0.22	0.15	0.11	0.72[a]	–0.088	0.238
Prior stroke[c]	1.37	0.27	0.68	0.49	0.842	11.042
Education level[d]	0.04	0.02	0.06	–0.04	0.152	–1.622
Smartphone ownership[c]	–0.07	0.09	0.21	–0.45	0.468	–8.295
Social media use[e]	0.37	–0.02	0.03	–0.06	0.030	–0.814
Cognitive impairment[f]	–0.05	–0.19	–0.17	0.16	–0.188	6.656
Depression[f]	0.37	0.14	0.47	–0.61	–0.158	–6.542
Anxiety[f]	0.21	0.44	0.17	–0.61	–0.158	–8.125

[a]Indicates statistical significance at $P<.05$.

[b]Unadjusted beta coefficients from univariate regression analysis.

[c]Coronary artery disease, coronary artery bypass graft, cardioversion, stroke, smartphone ownership are all coded as yes (1) or no (0).

[d]Education coded from 1 to 6, from completed some high school to some graduate degree.

[e]Social media use coded from 1 to 5, from no use to >6 hours per week.

[f]Cognitive impairment coded as Montreal Cognitive Assessment <26, depression coded as Patient Health Questionnaire–9 >4, anxiety coded as Generalized Anxiety Disorder–7 >4. Cognitive impairment, depression, and anxiety available for 22 participants.

Discussion

Principal Findings

We enrolled 40 older patients presenting to an ambulatory clinic in an observational study of mHealth devices for arrhythmia monitoring. Participants wore a smartwatch (the Samsung Simband 2.0, Figure 1) for about 40 min and followed a rigorous protocol to generate motion noise to simulate ADLs. Over 2000 30-second pulse segments were obtained, and after noise elimination, blinded analysis of over 300 30-second pulse segments was conducted using a real-time realizable algorithm. The pulse analysis algorithm demonstrated excellent sensitivity, specificity, and accuracy for the detection of an irregular pulse consistent with AF. Despite their advanced age, lack of familiarity with smartwatches, and high burden of cardiovascular and noncardiovascular comorbidities, study participants found the smartwatch highly acceptable overall and across several important usability domains. We did observe, however, that younger age and prior history of cardioversion were associated with greater comfort with use of the device for heart rhythm monitoring. Our findings demonstrate the accuracy and acceptability of a real-time realizable pulse-analysis algorithm analyzing PPG data from a smartwatch among participants at high risk for incident or recurrent AF.

Current Atrial Fibrillation Detection Modalities

Early AF detection results in treatment with oral anticoagulants that reduce stroke risk by up to 70% [6], and daily home monitoring for AF among stroke survivors reduces stroke risk by ~18% as compared with conventional in-office ECGs when obtained every 6 to 12 months [34]. In the randomized COMPAS trial, hospitalizations for atrial arrhythmias and stroke were higher in conventionally-monitored patients compared with those prescribed home monitoring, suggesting that home monitoring for AF has clinical impact and is cost effective [34]. Furthermore, in the STROKESTOP study [8], home monitoring using intermittent rhythm recordings identified a significant proportion of previously undiagnosed AF in an older cohort compared with usual care and has been shown in subsequent cost-effectiveness analyses to have a favorable quality-adjusted life years saved [35].

The most commonly prescribed noninvasive ECG monitors (24-hour Holter monitors) demonstrate low yield for AF detection across a wide array of high-risk subgroups, likely as the monitoring coverage of a 24- or 48-hour monitor is simply too brief [36]. For example, the median time to first AF episode was approximately 30 days in the Cryptogenic Stroke and Underlying Atrial Fibrillation study [16,37]. Long-term implantable monitoring significantly improves paroxysmal AF detection rates compared with conventional, shorter-term,

noninvasive monitors, but few patients elect to undergo this costly and invasive procedure [16].

Owing to the suspected prevalence and adverse health impact of undiagnosed AF, there is interest in developing new mHealth tools capable of enabling long-term, noninvasive monitoring. Recently, several companies have developed new software and hardware needed to harness the ubiquity and usability of commercial wrist-based wearable devices for AF screening and monitoring [10,12]. AliveCor (AliveCor Inc) recently released the KardiaBand, an FDA-approved, commercially available smartwatch wristband with an embedded ECG electrode designed to pair with the Apple Watch for AF monitoring [38]. The most recent version of this technology aims to use a PPG-based pulse irregularity notification feature to prompt the user to perform an active rhythm check. This is distinct from AliveCor's Kardia device and other commercial ECG-based devices, such as the MyDiagnostick [39] and Zenicor [40] products, which have no passive monitoring component. A recent study including 100 participants undergoing a cardioversion and wearing the AliveCor KardiaBand showed that, as in our cohort, MNA was common but AF identification was possible and fairly accurate compared with physician review on clean ECG segments [41,42]. Another recent study leveraging the Health eHeart study also demonstrated that an Apple Watch can also be used to collect pulse signals and discriminate AF with modest accuracy compared with self-reported AF using a deep learning neural network approach [19]. This approach has also been shown to be modestly accurate and is computationally intensive. Finally, the SmartWATCHes for Detection of Atrial Fibrillation study, which analyzed pulse data collected from smartwatches using different parameters of heart rate variability than this study (normalized RMSSD and ShE) also found that pulse data can be accurate for AF diagnosis [20]. These promising results, in conjunction with our own, further highlight the importance of addressing human factors such as usability and implementation considerations.

Like the Apple Watch–AliveCor KardiaBand dyad and similar to the recently FDA-cleared Apple Watch 4, the Samsung Simband 2.0 records PPG and single-lead ECG signals using an electrode embedded in the watch (Figure 1). The Samsung Simband platform offered several advantages, leading to its selection for use in our investigation. First, unlike Apple, Samsung provides investigators with unfettered access to a secure data management system designed for research (ARTIK Cloud). Second, Samsung allowed for independent development of open-source apps and data sharing between researchers.

Pulse-Based Atrial Fibrillation Detection

The pulse analysis approach tested in our study demonstrated high accuracy for the detection of AF using smartwatch pulse recordings despite the fact that participants were asked to perform activities intentionally designed to create MNAs [23,24]. Owing to the intensity of our motion protocols and vulnerability of wrist-based devices to signal corruption, our motion noise detection algorithm deemed 88% of the 30-second pulse segments to be corrupted. Prior investigators have observed similar rates of motion noise corruption when analyzing PPG data for heart rate analysis from smartwatches

[43]. The low coverage noted in our study represents a significant potential limitation of PPG-based technologies for heart rhythm monitoring, especially during waking hours and during active periods. This finding highlights the importance of noise detection and the need for algorithms to filter and address MNA [44,45], which may contribute to false positives (owing to noisy segments that are detected as having AF if the algorithms are not sufficiently robust) or false negatives (owing to segments that are eliminated because of noise but include AF if the algorithms are overly restrictive). Despite low rates of pulse coverage, we anticipate higher coverage rates in natural environments, especially during sedentary periods and during sleep [46]. From a diagnostic perspective, although 5 30-second pulse segments were incorrectly identified by the algorithm as being consistent with AF, the clinical significance of such brief episodes is unknown [47]. To date, only episodes lasting 5 min or more have been associated with risk for ischemic stroke [48]. We anticipate that future approaches using pulse data from a smartwatch will be tuned to match performance to clinical use [49], perhaps requiring multiple, sequential 30-second pulse segments to demonstrate pulse irregularity or confirmation with ECG, to reduce false positive AF detection and enhance usefulness.

Our approach to AF detection using a smartwatch involves passive pulse acquisition and real-time analysis using methods that are accurate but not computationally demanding [10,27]. Our approach enhances potential acceptability and successful implementation as it can be feasibly ported into a wide range of smartwatch devices with different hardware specifications, and will not to require any external devices.

Device Usability and Acceptability

Consistent with this hypothesis, most participants deemed the smartwatch system highly usable overall and expressed comfort with using the system for home heart rhythm monitoring. Our finding that the majority of participants owned a smartphone and were willing to use a smartwatch to monitor themselves debunks the commonly held misconceptions about older Americans at risk for AF and their facility with mobile devices, but is entirely consistent with an emerging literature showing that older Americans are increasing using smart devices and are open to using wearables for disease prevention and treatment [21]. Presently, over 85% of Americans over 65 years currently use mobile phones, a proportion that is increasing [50]. Smartphone use is becoming increasingly common among seniors (10% increase in the last 3 years) [50], and despite well recognized usability barriers, such as physical difficulties, skeptical attitudes, and difficulty learning new technologies, older users can adopt new technologies [51] and often prefer smartphones and smartwatches for running mHealth apps to conventional diagnostic devices, as was observed in our study (Figure 3) [52]. Our results also highlight the importance of proper patient and caregiver education when implementing novel mHealth interventions. Study staff were responsible for set up and deployment of the devices in each participant, which likely facilitated their user experience and contributed to the high perceived usability.

Although we did not find significant associations between any participant characteristics and overall smartwatch usability, we did observe that older age and history of coronary artery bypass graft surgery were associated with lesser comfort using the smartwatch for heart rhythm monitoring, consistent with prior studies [53]. We observed that participants with a history of prior cardioversion for AF were more comfortable using and sharing their heart rhythm data using the smartwatch (Table 2). This finding likely reflects the relative importance of AF monitoring among patients already affected by the disease and perhaps helps to identify an ideal *early adopter* population to test home use of smartwatches for heart rhythm monitoring.

In contrast to prior investigations, our study focused on establishing both the accuracy and usability of a scalable smartwatch-based approach to AF monitoring and screening among a population of older potential users during active periods intended to simulate ADLs. Not only will future studies need to be conducted to examine long-term adherence to smartwatches for AF monitoring among at-risk populations, further work to tune the AF detection algorithm for ideal performance using large, diverse study cohorts will be required. Our findings suggest, however, that older users, when provided support in learning to use smartwatches, can use them well, and that data derived from these devices are of sufficient quality so as to enable high quality rhythm analysis.

Strengths and Limitations

Our study has several strengths. First, in contrast to most mHealth studies involving younger participants, our study was able to enroll older participants at high risk for incident or recurrent AF. Second, in contrast to other mHealth studies that examine performance of signal processing algorithms from *ideal state* data, we employed a standardized protocol to introduce motion noise to simulate ADLs [23,24]. Third, we conducted blinded analysis of pulse data using validated methods and rigorous quality control [10,11,29]. Finally, despite the rapid development of new technologies for smartwatch-based AF monitoring and enthusiasm in the AF community, we are the first, to our knowledge, to explore usability and acceptability of smartwatches for cardiac monitoring in the elderly.

Several limitations of our study warrant presentation. Our sample was enrolled from an ambulatory clinic at a single tertiary care medical center in central Massachusetts and was relatively homogenous with regard to race and gender. This racial and gender homogeneity significantly limits our ability

to generalize these findings to members of other racial groups, community-dwelling individuals not under the care of a physician, or among individuals from other geographic areas. We plan to rectify this and enrich our study population by targeting a more diverse sample in future data collection. Furthermore, the sample was highly educated and reported a relatively high annual income, potentially limiting generalizability to less well-educated or poorer individuals. In addition, our sample was older, affected by a moderate to significant burden of physical and cognitive impairments, and the penetrance of smart device use was lower than the national average [50] (55% owned a smartphone), suggesting that our findings may underestimate the usability of the smartwatch for rhythm analysis in other groups. Our gold standard used was also based on a commercial automated algorithm, and although cardiologist overread performed on 20% of the sample matched the algorithm completely, manual overread on the remaining sample was not performed due to feasibility. In addition, although we intentionally designed study protocols to generate MNA from activities of normal daily life, we did not assess accuracy or usability with home use. The device was tested in a temperature- and light-controlled clinic environment, and disturbance of these factors, in addition to variables unmeasured in our study (such as skin turgor and humidity), may affect performance. Finally, our study population was a convenience sample enriched for participants with AF, and as such, the usability data collected may be affected by selection bias. This inflated AF prevalence also likely resulted in a higher positive predictive value for AF identification than in real-world settings, where the AF prevalence is much lower.

Conclusions

A novel, real-time realizable software algorithm analyzing pulse data from a smartwatch exhibits excellent performance for the detection of an irregular pulse consistent with AF among older individuals creating motion noise to simulate ADLs. Furthermore, the smartwatch system was deemed highly usable by older participants enrolled in our study, suggesting that long-term monitoring for AF using wrist-based mHealth devices holds promise. Future work is needed to assess provider impressions of the system, to validate findings from our study in much larger and more diverse cohorts, and examine long-term adherence to daily home use, as well as in-field accuracy of AF diagnosis among older individuals at risk for incident or recurrent AF.

presubmission, but the Samsung team had no input into the design or analysis plan, and the content in the paper was solely generated by the research team. This project was directly supported by the NIH grant 1R01HL137734.

Authors' Contributions

ED was responsible for data analysis/interpretation, drafting majority of the paper, statistics, and data collection. DH, CW, SKB, and OA were responsible for data analysis/interpretation, drafting sections of the paper, critical revision of the paper, statistics, and data collection. AS, JS, TF, MM, and SL were responsible for data interpretation and critical revision of the paper. DL and BB were responsible for data analysis and statistics. HTH was responsible for critical revision of article and approval of the paper. KC and DDM were responsible for concept/design, drafting, critical revision and approval of the paper, and securing of grant funding.

Conflicts of Interest

Multimedia Appendix 1
Descriptive statistics of participant responses to smartwatch usability assessment.
[DOCX File, 13KB - cardio_v3i1e13850_app1.docx]

References

1. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 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 Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol 2014 Dec 2;64(21):e1-76 [FREE Full text] [doi: 10.1016/j.jacc.2014.03.022] [Medline: 24685669]

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. J Am Med Assoc 2001 May 9;285(18):2370-2375. [Medline: 11343485]

3. Savelieva I, Camm AJ. Clinical relevance of silent atrial fibrillation: prevalence, prognosis, quality of life, and management. J Interv Card Electrophysiol 2000 Jun;4(2):369-382. [Medline: 10936003]

4. Lin HJ, Wolf PA, Kelly-Hayes M, Beiser AS, Kase CS, Benjamin EJ, et al. Stroke severity in atrial fibrillation. The Framingham Study. Stroke 1996 Oct;27(10):1760-1764. [Medline: 8841325]

5. Friberg L, Hammar N, Pettersson H, Rosenqvist M. Increased mortality in paroxysmal atrial fibrillation: report from the Stockholm Cohort-Study of Atrial Fibrillation (SCAF). Eur Heart J 2007 Oct;28(19):2346-2353. [doi: 10.1093/eurheartj/ehm308] [Medline: 17670754]

6. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007 Jun 19;146(12):857-867. [Medline: 17577005]

7. Jaakkola J, Mustonen P, Kiviniemi T, Hartikainen J, Palomäki A, Hartikainen P, et al. Stroke as the first manifestation of atrial fibrillation. PLoS One 2016;11(12):e0168010 [FREE Full text] [doi: 10.1371/journal.pone.0168010] [Medline: 27936187]

8. Svennberg E, Engdahl J, Al-Khalili F, Friberg L, Frykman V, Rosenqvist M. Mass screening for untreated atrial fibrillation: the STROKESTOP Study. Circulation 2015 Jun 23;131(25):2176-2184 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.114.014343] [Medline: 25910800]

9. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016 Oct 7;37(38):2893-2962. [doi: 10.1093/eurheartj/ehw210] [Medline: 27567408]

10. McManus DD, Lee J, Maitas O, Esa N, Pidikiti R, Carlucci A, et al. A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. Heart Rhythm 2013 Mar;10(3):315-319 [FREE Full text] [doi: 10.1016/j.hrthm.2012.12.001] [Medline: 23220686]

11. McManus DD, Chong JW, Soni A, Saczynski JS, Esa N, Napolitano C, et al. PULSE‐SMART: pulse‐based arrhythmia discrimination using a novel smartphone application. J Cardiovasc Electrophysiol 2016 Jan;27(1):51-57. [doi: 10.1111/jce.12842] [Medline: 26391728]

12. Guo Y, Chen Y, Lane DA, Liu L, Wang Y, Lip GY. Mobile health technology for atrial fibrillation management integrating decision support, education, and patient involvement: mAF App Trial. Am J Med 2017 Dec;130(12):1388-96.e6 [FREE Full text] [doi: 10.1016/j.amjmed.2017.07.003] [Medline: 28847546]

13. Lin L, Bessette A. Apple Newsroom. Redesigned Apple Watch Series 4 revolutionizes communication, fitness and health URL: https://www.apple.com/newsroom/2018/09/redesigned-apple-watch-series-4-revolutionizes-communication-fitness-and-health/ [accessed 2019-02-20] [WebCite Cache ID 76M5J1ttT]

14. Medical College of Wisconsin. Local doctor onstage at event and discussed new health functionality of Apple Watch Series 4 URL: https://www.mcw.edu/newsroom/news-articles/local-doctor-onstage-at-event-and-discussed-new-health-functionality-of-apple-watch-series-4 [accessed 2019-02-21] [WebCite Cache ID 76M5mrUke]

15. Halcox JPJ, Wareham K, Cardew A, Gilmore M, Barry JP, Phillips C, et al. Assessment of remote heart rhythm sampling using the AliveCor heart monitor to screen for atrial fibrillation: the REHEARSE-AF study. Circulation 2017 Nov 7;136(19):1784-1794. [doi: 10.1161/CIRCULATIONAHA.117.030583] [Medline: 28851729]

16. Sanna T, Diener H, Passman RS, Di LV, Bernstein RA, Morillo CA, et al. Cryptogenic stroke and underlying atrial fibrillation. N Engl J Med 2014 Jun 26;370(26):2478-2486. [doi: 10.1056/NEJMoa1313600] [Medline: 24963567]

17. Carpenter A, Frontera A. Smart-watches: a potential challenger to the implantable loop recorder? Europace 2016 Jun;18(6):791-793. [doi: 10.1093/europace/euv427] [Medline: 26847074]

18. Leader N, Dorian P, Lam J, Lee C, Woo A, Chow C. Evaluation of heart rate trackers in patients with atrial fibrillation. Can J Cardiol 2018 Oct;34(10):S152-S153. [doi: 10.1016/j.cjca.2018.07.174]

19. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. JAMA Cardiol 2018 May 1;3(5):409-416. [doi: 10.1001/jamacardio.2018.0136] [Medline: 29562087]

20. Dörr M, Nohturfft V, Brasier N, Bosshard E, Djurdjevic A, Gross S, et al. The WATCH AF trial: SmartWATCHes for detection of atrial fibrillation. JACC Clin Electrophysiol 2019 Feb;5(2):199-208. [doi: 10.1016/j.jacep.2018.10.006] [Medline: 30784691]

21. Kruse CS, Mileski M, Moreno J. Mobile health solutions for the aging population: a systematic narrative analysis. J Telemed Telecare 2017 May;23(4):439-451. [doi: 10.1177/1357633X16649790] [Medline: 27255207]

22. Dao D, Salehizadeh SMA, Noh Y, Chong JW, Cho CH, McManus D, et al. A robust motion artifact detection algorithm for accurate detection of heart rates from photoplethysmographic signals using time–frequency spectral features. IEEE J Biomed Health Inform 2017 Dec;21(5):1242-1253. [doi: 10.1109/JBHI.2016.2612059] [Medline: 28113791]

23. Chong JW, Dao DK, Salehizadeh SM, McManus DD, Darling CE, Chon KH, et al. Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection-reduction approach. Part I: Motion and noise artifact detection. Ann Biomed Eng 2014 Nov;42(11):2238-2250. [doi: 10.1007/s10439-014-1080-y] [Medline: 25092422]

24. Salehizadeh SM, Dao DK, Chong JW, McManus D, Darling C, Mendelson Y, et al. Photoplethysmograph signal reconstruction based on a novel motion artifact detection-reduction approach. Part II: Motion and noise artifact removal. Ann Biomed Eng 2014 Nov;42(11):2251-2263. [doi: 10.1007/s10439-014-1030-8] [Medline: 24823655]

25. Samsung Simband. Simband Applications URL: https://www.simband.io/documentation/simband-documentation/applications/ [accessed 2019-02-20] [WebCite Cache ID 76M61o2Xd]

26. Bashar S, Han D, Soni A, McManus D, Chon K. Developing a novel noise artifact detection algorithm for smartphone PPG signals: Preliminary results. 2018 Apr 9 Presented at: 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI); March 4-7, 2018; Las Vegas, NV p. 79. [doi: 10.1109/BHI.2018.8333374]

27. Dash S, Chon KH, Lu S, Raeder EA. Automatic real time detection of atrial fibrillation. Ann Biomed Eng 2009 Sep;37(9):1701-1709. [doi: 10.1007/s10439-009-9740-z] [Medline: 19533358]

28. Harrington JL, Chong JW, Li J, Esa N, Pidikiti R, Maitas O, et al. The detection and differentiation of arrrhythmias using a smartphone: A clinical study of patients with atrial fibrillation, premature atrial and premature ventricular contractions. J Am Coll Cardiol 2013 Mar;61(10):E362. [doi: 10.1016/S0735-1097(13)60362-9]

29. Lee J, Nam Y, McManus DD, Chon KH. Time-varying coherence function for atrial fibrillation detection. IEEE Trans Biomed Eng 2013 Oct;60(10):2783-2793. [doi: 10.1109/TBME.2013.2264721] [Medline: 23708769]

30. Brooke J. SUS-a quick and dirty usability scale. In: Usability Evaluation In Industry. London: Taylor & Francis; 1996:189-194.

31. Freitas S, Simões MR, Alves L, Santana I. Montreal cognitive assessment: validation study for mild cognitive impairment and Alzheimer disease. Alzheimer Dis Assoc Disord 2013;27(1):37-43. [doi: 10.1097/WAD.0b013e3182420bfe] [Medline: 22193353]

32. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]

33. Spitzer RL, Kroenke K, Williams JB, 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]

34. Mabo P, Victor F, Bazin P, Ahres S, Babuty D, da Costa A, et al. A randomized trial of long-term remote monitoring of pacemaker recipients (the COMPAS trial). Eur Heart J 2012 May;33(9):1105-1111 [FREE Full text] [doi: 10.1093/eurheartj/ehr419] [Medline: 22127418]

XSL·FO
RenderX

35. Aronsson M, Svennberg E, Rosenqvist M, Engdahl J, Al-Khalili F, Friberg L, et al. Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ECG recording. Europace 2015 Jul;17(7):1023-1029. [doi: 10.1093/europace/euv083] [Medline: 25868469]

36. Ziegler PD, Koehler JL, Mehra R. Comparison of continuous versus intermittent monitoring of atrial arrhythmias. Heart Rhythm 2006 Dec;3(12):1445-1452. [doi: 10.1016/j.hrthm.2006.07.030] [Medline: 17161787]

37. Gladstone DJ, Spring M, Dorian P, Panzov V, Thorpe KE, Hall J, EMBRACE Investigators and Coordinators. Atrial fibrillation in patients with cryptogenic stroke. N Engl J Med 2014 Jun 26;370(26):2467-2477. [doi: 10.1056/NEJMoa1311376] [Medline: 24963566]

38. KardiaBand. AliveCor Store URL: https://store.alivecor.com/products/kardiaband [accessed 2018-02-20] [WebCite Cache ID 76MBaq4XY]

39. Vaes B, Stalpaert S, Tavernier K, Thaels B, Lapeire D, Mullens W, et al. The diagnostic accuracy of the MyDiagnostick to detect atrial fibrillation in primary care. BMC Fam Pract 2014 Jun 9;15:113 [FREE Full text] [doi: 10.1186/1471-2296-15-113] [Medline: 24913608]

40. Doliwa PS, Frykman V, Rosenqvist M. Short-term ECG for out of hospital detection of silent atrial fibrillation episodes. Scand Cardiovasc J 2009 Jun;43(3):163-168. [doi: 10.1080/14017430802593435] [Medline: 19096977]

41. Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, et al. Smartwatch algorithm for automated detection of atrial fibrillation. J Am Coll Cardiol 2018 May 29;71(21):2381-2388. [doi: 10.1016/j.jacc.2018.03.003] [Medline: 29535065]

42. Chon KH, McManus DD. Detection of atrial fibrillation using a smartwatch. Nat Rev Cardiol 2018 Nov;15(11):657-658. [doi: 10.1038/s41569-018-0057-1] [Medline: 29985453]

43. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. J Sports Sci Med 2016 Sep;15(3):540-547 [FREE Full text] [Medline: 27803634]

44. Hoog Antink C, Schulz F, Leonhardt S, Walter M. Motion artifact quantification and sensor fusion for unobtrusive health monitoring. Sensors (Basel) 2017 Dec 25;18(1):38 [FREE Full text] [doi: 10.3390/s18010038] [Medline: 29295594]

45. Bhowmik T, Dey J, Tiwari V. A novel method for accurate estimation of HRV from smartwatch PPG signals. 2017 39th Annu Int Conf IEEE Eng Med Biol Soc EMBC Internet Seogwipo: IEEE; 2017 Dec Presented at: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 11-15, 2017; Seogwipo, South Korea p. 109-112. [doi: 10.1109/EMBC.2017.8036774]

46. Dooley E, Golaszewski N, Bartholomew J. Estimating accuracy at exercise intensities: a comparative study of self-monitoring heart rate and physical activity wearable devices. JMIR Mhealth Uhealth 2017 Mar 16;5(3):e34 [FREE Full text] [doi: 10.2196/mhealth.7043] [Medline: 28302596]

47. Swiryn S, Orlov MV, Benditt DG, DiMarco JP, Lloyd-Jones DM, Karst E, RATE Registry Investigators. Clinical implications of brief device-detected atrial tachyarrhythmias in a cardiac rhythm management device population: results from the registry of atrial tachycardia and atrial fibrillation episodes. Circulation 2016 Dec 18;134(16):1130-1140. [doi: 10.1161/CIRCULATIONAHA.115.020252] [Medline: 27754946]

48. Glotzer TV, Hellkamp AS, Zimmerman J, Sweeney MO, Yee R, Marinchak R, MOST Investigators. Atrial high rate episodes detected by pacemaker diagnostics predict death and stroke: report of the Atrial Diagnostics Ancillary Study of the MOde Selection Trial (MOST). Circulation 2003 Apr 1;107(12):1614-1619. [doi: 10.1161/01.CIR.0000057981.70380.45] [Medline: 12668495]

49. Koshy AN, Sajeev JK, Nerlekar N, Brown AJ, Rajakariar K, Zureik M, et al. Utility of photoplethysmography for heart rate estimation among inpatients. Intern Med J 2018 Dec;48(5):587-591. [doi: 10.1111/imj.13777] [Medline: 29722189]

50. Anderson M, Perrin A. Pew Research Center. Technology use among seniors Internet URL: http://www.pewinternet.org/2017/05/17/technology-use-among-seniors/ [accessed 2018-03-07] [WebCite Cache ID 76MCU1vnu]

51. Levine DM, Lipsitz SR, Linder JA. Trends in seniors' use of digital health technology in the United States, 2011-2014. J Am Med Assoc 2016 Aug 2;316(5):538-540. [doi: 10.1001/jama.2016.9124] [Medline: 27483069]

52. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]

53. Tully PJ, Baker RA. Depression, anxiety, and cardiac morbidity outcomes after coronary artery bypass surgery: a contemporary and practical review. J Geriatr Cardiol 2012 Jun;9(2):197-208 [FREE Full text] [doi: 10.3724/SP.J.1263.2011.12221] [Medline: 22916068]

Abbreviations

ADLs: activities of daily living
AF: atrial fibrillation
AHA: American Heart Association
ECG: electrocardiogram
EHR: electronic health record

FDA: Food and Drug Administration
ID: identification number
mHealth: mobile health
MNA: motion noise artifact
NIH: National Institute of Health
NSR: normal sinus rhythm
PAC: premature atrial contraction
PPG: photoplethysmogram
PVC: premature ventricular contraction
RMSSD: root mean square of successive difference of RR intervals
SampEn: sample entropy
ShE: Shannon Entropy
SUS: System Usability Scale
UConn: University of Connecticut
UMMS: University of Massachusetts Medical School

XSL•FO
RenderX

Original Paper

Use of Free-Living Step Count Monitoring for Heart Failure Functional Classification: Validation Study

Jonathan-F Baril[1,2], MHSc; Simon Bromberg[1,2], MHSc; Yasbanoo Moayedi[3], MD; Babak Taati[2,4], PhD, PEng; Cedric Manlhiot[3], PhD; Heather Joan Ross[2,3], MD, MHSc; Joseph Cafazzo[1,2,5], PhD, PEng

[1]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

[2]Institute of Biomaterials & Biomedical Engineering, University of Toronto, Toronto, ON, Canada

[3]Ted Rogers Centre of Excellence in Heart Function, University Health Network, Toronto, ON, Canada

[4]Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

[5]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Abstract

Background: The New York Heart Association (NYHA) functional classification system has poor inter-rater reproducibility. A previously published pilot study showed a statistically significant difference between the daily step counts of heart failure (with reduced ejection fraction) patients classified as NYHA functional class II and III as measured by wrist-worn activity monitors. However, the study's small sample size severely limits scientific confidence in the generalizability of this finding to a larger heart failure (HF) population.

Objective: This study aimed to validate the pilot study on a larger sample of patients with HF with reduced ejection fraction (HFrEF) and attempt to characterize the step count distribution to gain insight into a more objective method of assessing NYHA functional class.

Methods: We repeated the analysis performed during the pilot study on an independently recorded dataset comprising a total of 50 patients with HFrEF (35 NYHA II and 15 NYHA III) patients. Participants were monitored for step count with a Fitbit Flex for a period of 2 weeks in a free-living environment.

Results: Comparing group medians, patients exhibiting NYHA class III symptoms had significantly lower recorded 2-week mean daily total step count (3541 vs 5729 [steps], *P*=.04), lower 2-week maximum daily total step count (10,792 vs 5904 [steps], *P*=.03), lower 2-week recorded mean daily mean step count (4.0 vs 2.5 [steps/minute], *P*=.04,), and lower 2-week mean and 2-week maximum daily per minute step count maximums (88.1 vs 96.1 and 111.0 vs 123.0 [steps/minute]; *P*=.02 and .004, respectively).

Conclusions: Patients with NYHA II and III symptoms differed significantly by various aggregate measures of free-living step count including the (1) mean and (2) maximum daily total step count as well as by the (3) mean of daily mean step count and by the (4) mean and (5) maximum of the daily per minute step count maximum. These findings affirm that the degree of exercise intolerance of NYHA II and III patients as a group is quantifiable in a replicable manner. This is a novel and promising finding that suggests the existence of a possible, completely objective measure of assessing HF functional class, something which would be a great boon in the continuing quest to improve patient outcomes for this burdensome and costly disease.

KEYWORDS

Introduction

Heart failure (HF), a global epidemic [1,2], is a complex chronic progressive condition associated with significant morbidity and mortality. HF is the leading cause of hospitalizations in the country, costing Canadians an estimated Can \$3 billion annually [3]. From both a systems and patient-centered perspective, clinicians caring for patients with HF have a strong desire to

reduce hospitalizations [3,4]. To do so, it is important for clinicians to be able to reliably assess disease progression and severity.

One of the ways in which HF is categorized is by the degree to which a patient's left ventricle retains the ability to pump out the blood it receives—known as the left ventricular ejection fraction (LVEF) [5,6]. The degree to which ejection fraction (EF) is reduced can be an indicator of what part of, and to what

degree, the heart has been damaged [5]. Practice guidelines recommend different interventional strategies according to the degree of preserved (or reduced) EF [5]. Broadly speaking, patients with an LVEF ≤40% are classified as suffering from a subtype of HF known as HF with reduced EF (HFrEF) [5,6]. Those with preserved EF are labeled as suffering from HF with preserved EF (HFpEF). Both subtypes are fairly common, with HFpEF comprising approximately 44% to 72% of cases, although it is difficult to make precise estimates as the exact LVEF cut-off for HFpEF versus HFrEF has varied over time and across geographic regions [6]. Nevertheless, current estimates indicate that HFpEF is starting to emerge as the most prevalent HF subtype (compared with HFrEF) in Canada and the United States, especially relative to the rest of the world [6,7].

Although decidedly more common in patients with HFrEF, the primary cause of HF overall is most commonly attributable to coronary heart disease (CHD): about 23% to 73% of patient cases depending on the study in question [8]. Hypertension (HT), often more associated with patients suffering from HFpEF, follows second as the hierarchy of competing common etiologies; of course, both CHD and HT commonly coexist in the same patient, which makes identifying the causal primacy of each condition difficult, especially as both CHD and HT are known to cause either type of HF [5,8]. For example, an analysis of patients in the well-known Framingham Heart Study showed that 63% of the 314 patients with HFrEF had CHD identified as the primary cause compared to 19% with HT identified as the primary cause [9]. In contrast, of the 220 patients with HFpEF, only 37% had CHD identified as the primary cause versus 36% with HT as the primary cause [9]. Of course, HF has many other known causes including valvular disease, congenital cardiac malformations, and pathogenic, nutritional, or toxicological causes, but CHD and HT are by far the most common [5].

As a result of the etiology of HF, in Canada, although not exclusively a disease of old age, HF prevalence and incidence increases sharply among Canadians aged 65 years and older, as expected from the high incidence and prevalence of cardiovascular disease (and CHD and HT in particular) among this subpopulation [4,10,11]. According to the Canadian Chronic Disease Surveillance System, in 2015, the crude prevalence rates among Canadian men aged between 40 and 49, 50 and 64, 65 and 79, and 80+ years were 0.34, 1.82, 7.07, and 20.02 (%), respectively, with slightly lower prevalence rates among women of the same age brackets at 0.23, 1.07, 4.65, and 17.92 (%), respectively [11]. In the last decade and a half of reported data (2000 to 2015), the age-standardized prevalence (among those aged 40 years and older) has also remained fairly constant, hovering around a mean (SD) of 3.07 % (SD 0.10 %^{2}) for women and approximately 31% higher for men at 4.03 % (SD 0.09 %^{2}) [11]. The incidence rate (for the same subpopulation), however, declined over the same period, from a peak of 952 to

612 (per 100,000) for men and from a peak of 714 to 459 (per 100,000) for women [11]. No data were recorded for those aged younger than 40 years [11].

One of the main manifestations of HF across populations is exercise intolerance [5,12]. As a result, apart from evaluating LVEF (among other biometrics), evaluating exercise intolerance forms an integral part of HF care and also constitutes an important widely used prognostic marker [12]. The New York Heart Association (NYHA) classification system is a formal system for assessing the functional exercise capacity of a patient where a higher NYHA class (IV, III) is associated with an increase in experienced HF symptoms, a decreased quality of life, and poor survival [13-15]. This classification system is highly subjective [12,16], with low inter-rater reliability, especially for NYHA class II and III [17]. The application of the criteria, thus, varies widely based on the patients' self-report and the individual physician's interpretation [12,16]. A quantifiable measure that removes this subjectivity to make the assessment of NYHA class more repeatable and objective would be beneficial.

A previous exploratory study [18] investigated wearable activity trackers in patients with HF and demonstrated a statistically significant difference between the daily average step counts (a proxy for exercise intolerance) in patients exhibiting NYHA class II and III symptoms. However, the study's small sample (n=8) limits scientific confidence in the generalizability of these findings. The primary aim of this study was to determine if these findings can be replicated using a larger sample collected independently from the original pilot study data. A secondary aim was to investigate wearable activity tracker usage by patients with HF and begin to characterize the step count distribution of these patients under free-living conditions in hopes of enabling the engineering of objective methods of assessing and monitoring NYHA functional class and, thereby, improving the ability of clinicians to accurately assess disease progression and severity.

Methods

Ethics Approval

This study is covered by institutional and research ethics approval (REB #14-7595) received from the University Health Network REB; (signed) informed consent was obtained from all study participants.

Recruitment

Patients in a moderately larger dataset (n=50) were consecutively recruited, as part of a broader umbrella study, from the Heart Function Clinic at Toronto General Hospital (TGH) in Toronto, Canada, from September 2014 to June 2015. The inclusion and exclusion criteria used are outlined in Textbox 1.

Textbox 1. Study inclusion and exclusion criteria.

Inclusion criteria:

- Adults (aged older than 18 years)

- Stable chronic heart failure

- New York Heart Association class II or III

- Left ventricular ejection fraction ≤35% (arising out of research requirements for the broader umbrella study)

- Able to walk without walking aids

- Capable of undergoing consent, understanding English instructions, and complying with the use of the study devices.

Exclusion criteria:

- Congenital heart disease

- Diagnosis less than 6 months before recruitment

- Traveling out of Canada for more than 1 week during the study period (to limit study costs–ie roaming charges)

Data Collection

Patients were supplied with a Fitbit Flex [19], an Android smartphone (Moto-G), the associated charging equipment for both devices, as well as a cellular internet data plan to facilitate syncing the tracker to the Fitbit server. Patients were instructed to wear the Fitbit daily on the same wrist, preferably their nondominant hand, for a period of 2 weeks, except during water activities such as showering or swimming, as the Flex is not waterproof. Patients were also instructed to charge the Fitbit at least every 3 days, preferably while they slept. The Fitbit data were retrieved using an open-source script published and available on GitHub and adapted for this study [20].

Population

Patients in our larger dataset were labeled as either NYHA class II and III or (according to standard practice in our clinic) when a physician was uncertain about the classification or felt that patients exhibited symptoms from different class levels, as a borderline or mixed class: I/II or II/III. As NYHA class I/II and II/III are not formally recognized NYHA classes, to perform our analysis, the authors regrouped borderline patients into one of the traditional 4 class NYHA according to the most extreme NYHA class in the mix and according to the following rationale: as NYHA class I corresponds to "no limitation of physical activity," [15] an absolute binary (yes/no) distinction, a patient assigned as class I/II, who necessarily must be exhibiting strictly more than "no limitation of physical activity" [15] (however slight) can be reasonably grouped with class II patients generally (those exhibiting "a slight limitation of physical activity") [15]. We designated this class I/II and class II group as NYHA group II*.

We extended the same line of reasoning for II/III patients, noting that patients assigned as class II/III must have experienced some *more* marked limitation of physical activity beyond that seen in patients classified in class II. As such, for consistency, we

grouped them with the lower class III. We designated this class II/III and III group as NYHA group III*.

Statistics

Consistent with our previous study [18], we used the Kruskal-Wallis rank test to compare the experimental variables of interest, including the mean daily total step count. As the data are clearly not normally distributed—as can be seen in Figure 1—and in keeping with the secondary aim of the study, we also computed various additional statistical summaries of the minute-by-minute step count data to attempt to better characterize the data distribution. To calculate these summaries, we performed a first aggregation: calculating statistical summaries (mean, SD; 5-number summaries; interquartile range [IQR]; skewness; and kurtosis) across each patient's individual patient-day of step data and then a second aggregation across the day summaries, calculating the max, min, mean, and SD of each patient's daily summaries for the 2-week period (producing a maximum of mean daily step counts, minimum of mean daily step counts, and mean of mean daily step counts) to assess overall variation across patient-days. The methodology is shown graphically in Figure 2. In addition, we generated statistical summaries treating the overall 2-week period as 1 continuous time period (instead of analyzing it day-by-day) and simply performed a single (1st) aggregation over that period to generate the corresponding statistical summary for that patient-period. We then performed a Kruskal-Wallis rank test on each of the generated statistical summaries and reported the corresponding median value of each NYHA group and the calculated unadjusted P value from the statistical test. Note that as we report unadjusted P values (ie, without multigroup correction), statistical significance should be interpreted in light of this limitation; rejection of the null hypothesis (ie, rejecting group II* statistical summary X=group III* statistical summary X) is, therefore, limited to that statistical summary alone—that is, in isolation from the other statistical tests performed. The analysis was performed using R [21], RStudio [22] with supporting packages [23-28].

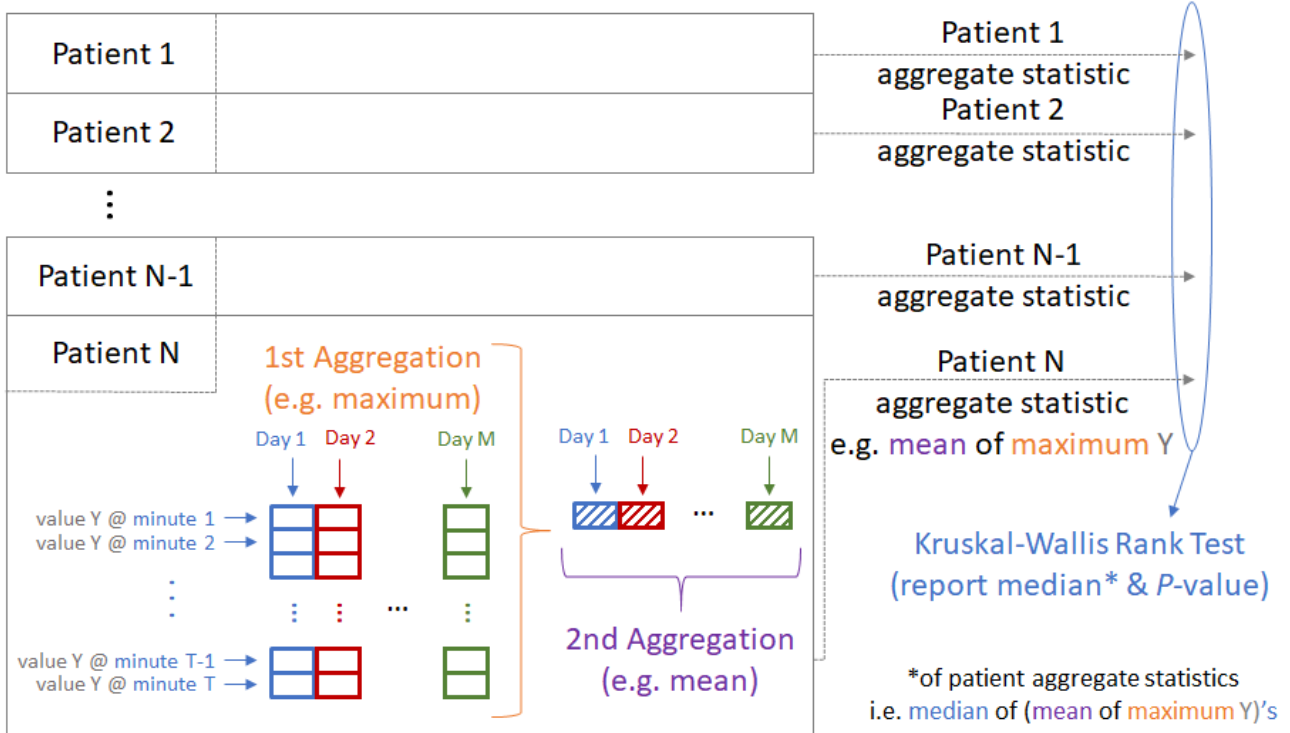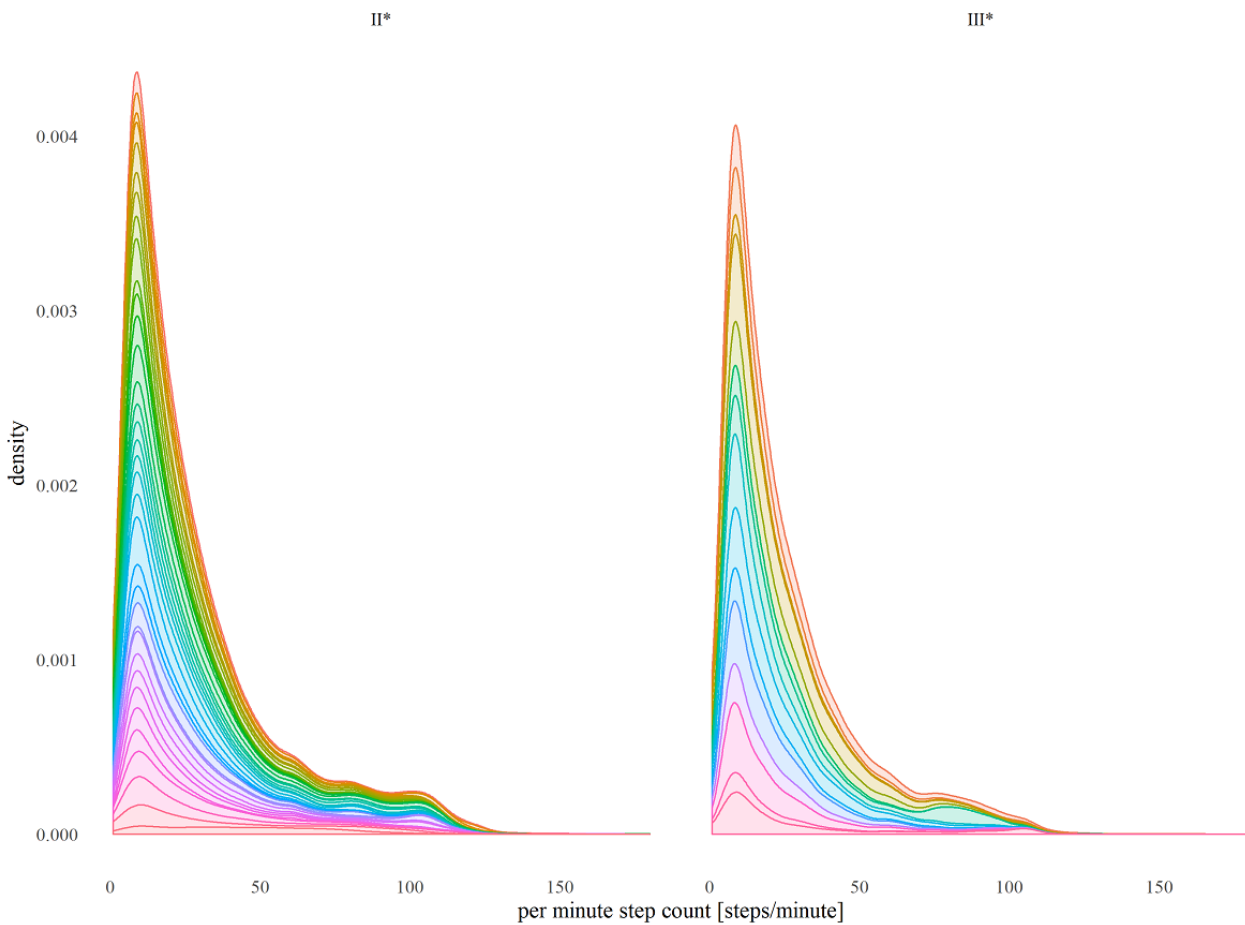**Figure 1.** Summary Statistic Computation Methodology.

Figure 2. Combined (all patients) distribution of per minute step counts by NYHA group (only step count values > 0). Colored internal segments illustrate relative contributions to distribution by each study participant.

Results

Table 1 provides demographic information for each of the patients in the dataset according to their NYHA class. Table 2 provides demographic information for the overall dataset and for patients when the dataset is regrouped according to the labeling scheme described in the Methods section (Population subsection). The patients are predominantly male (83% vs 93%), aged (median [IQR]): 55 (19) vs 56 (18) years, and overweight (body mass index (median [IQR]): 27.1 (7.6) vs 29.6 (6.6) kg/m^2).

Table 3 includes results that were found to be significant at the P=.05 level of significance (reported as median values because of the use of the nonparametric Kruskal-Wallis rank test). Table 4 contains the remaining nonsignificant results excluding any statistical summary that returned a 0 value for all classes (eg, aggregations involving daily or overall minimum, 1st, 2nd, and 3rd quartile) because of the overwhelming frequency of 0 per minute step count. The mean daily total steps and the mean and max of daily per minute step count maximums are plotted graphically in Figures 3, 4, and 5, respectively.

Table 1. Study dataset demographics (by original New York Heart Association class label).

Variable	NYHA[a] I/II	NYHA II	NYHA II/III	NYHA III
Participants, n (%)	9 (18)	26 (52)	4 (8)	11 (22)
Number of males, n (%)	6 (67)	23 (89)	4 (100)	10 (91)
Age (years), Q1[b]|M[c]|Q3[d]	50|52|62	45|57|66	45|50|56	53|58|68
Height (cm), Q1|M|Q3	167|172|180	167|172|180	167|172|180	167|172|180
Weight (kg), Q1|M|Q3	60.0|84.8|96.0	79.0|84.5|93.8	80.8|96.2|103.8	82.0|94.0|104.0
BMI (kg/m^2), Q1|M|Q3	21.5|24.0|29.3	25.0|27.6|31.7	25.8|30.4|33.0	27.0|29.6|32.8

[a]NYHA: New York Heart Association.

[b]Q1: 1st quartile.

[c]M: median.

[d]Q3: 3rd quartile.

Table 2. Study regrouped dataset demographics (Overall, New York Heart Association group II[*] and III[*]).

Variable	Overall	NYHA[a] Group II[*]	NYHA Group III[*]
Total participants, n (%)	50 (100)	35 (70)	15 (30)
Number of males, n (%)	43 (86)	29 (83)	14 (93)
Age (years), Q1[b]|M[c]|Q3[d]	47|55|65	46|55|65	49|56|67
Height (cm), Q1|M|Q3	170|175|180	168|175|179	171|177|180
Weight (kg), Q1|M|Q3	74.9|89.0|96.5	73.4|84.8|95.0	82.0|94.0|104.2
BMI (kg/m^2), Q1|M|Q3	24.7|28.1|32.1	24.0|27.1|31.6	27.0|29.6|33.6

[a]NYHA: New York Heart Association.

[b]Q1: 1st quartile.

[c]M: median.

[d]Q3: 3rd quartile.

Table 3. Significant findings for comparisons between group II* and group III*.

Variable	Group II[*] (=I/II+II), median	Group III[*] (=II/III+III), median	P value
Maximum			
Maximum 2-week PMSC[a] (steps/minute)	123.0	111.0	.004[b]
Maximum of maximum DPMSC[c] (steps/minute)	123.0	111.0	.004[b]
Mean of maximum DPMSC (steps/minute)	96.1	88.1	.02[d]
Mean			
Mean 2-week PMSC (steps/minute)	4.0	2.5	.04[d]
Maximum of mean DPMSC (steps/minute)	7.5	4.1	.03[d]
Mean of mean DPMSC (steps/minute)	4.0	2.5	.04[d]
SD of mean DPMSC ($steps^2/minute^2$)	1.8	1.1	.04[d]
SD			
SD of 2-week PMSC ($steps^2/minute^2$)	13.3	9.2	.02[d]
Maximum of DPMSC SD ($steps^2/minute^2$)	20.6	14.5	.002[b]
Mean of DPMSC SD ($steps^2/minute^2$)	12.0	8.8	.03[d]
Total			
Total 2-week SC[e] (steps)	88130	53123	.03[d]
Maximum of total DPMSC (steps)	10792	5904	.03[d]
Mean of total DPMSC (steps)	5729	3541	.04[d]
SD of total DPMSC ($steps^2$)	2570	1513	.04[d]

[a]PMSC: per minute step count.

[b]$P<.01$.

[c]DPMSC: daily per minute step count.

[d]$P<.05$.

[e]SC: step count.

XSL•FO

RenderX

Table 4. Nonsignificant findings for comparisons between group II* and group III*.

Variable	Group II* (=I/II+II), median	Group III* (=II/III+III), median	P value
Demographics			
Sex (male=0, female=1)	0	0	.33
Age (years)	55	56	.71
Height (cm)	175.0	177.0	.38
Weight (kg)	84.8	94.0	.17
BMI[a] (kg/m^2)	27.1	29.6	.28
Righthanded?[b] (no=0, yes=1)	1	1	.18
Wristband preference[c] (left=0, right=1)	0	0	.16
Maximum			
SD of maximum DPMSC[d] (steps^2/minute^2)	24.6	23.5	.76
Minimum of maximum DPMSC (steps/minute)	42.5	34.7	.58
75th percentile			
Maximum of 75th percentile of DPMSC (steps/minute)	0	0	.93
Mean of 75th percentile of DPMSC (steps/minute)	0	0	.89
SD of 75th percentile of DPMSC (steps/minute)	0	0	.91
Mean			
Minimum of mean DPMSC (steps/minute)	0.3	0.1	.90
Median			
Median of 2-week PMSC[e] (steps/minute)	0	0	N/A
Maximum of median DPMSC (steps/minute)	0	0	N/A
Minimum of median DPMSC (steps/minute)	0	0	N/A
Total			
Minimum of total DPMSC (steps)	420	164	.90
Interquartile range (IQR)			
Maximum of DPMSC IQR (steps/minute)	0	0	.93
Mean of DPMSC IQR (steps/minute)	0	0	.89
SD of DPMSC IQR (steps^2/minute$^{2)}$	0	0	.91
SD			
Minimum of DPMSC SD (steps^2/minute^2)	2.9	1.2	.80
Skewness			
2-week PMSC skewness	4.6	5.5	.29
Maximum of daily SC[f] skewness	8.8	8.5	.97
Mean of daily SC skewness	4.9	5.1	.76
SD of daily SC skewness	1.3	1.4	.76
Minimum of daily SC skewness	3.3	3.4	.65
Kurtosis			
2-week PMSC kurtosis	24.5	36.0	.25
Maximum of daily SC kurtosis	99.3	99.4	.97
Mean of daily SC kurtosis	31.7	33.4	.71
SD of daily SC kurtosis	20.1	22.8	.73

XSL•FO

RenderX

Variable	Group II[*] (=I/II+II), median	Group III[*] (=II/III+III), median	P value
Minimum of daily SC kurtosis	10.4	13.2	.47

[a]BMI: body mass index.

[b]Is patient righthanded?

[c]Right- or lefthanded preference for wristband.

[d]DPMSC: daily per minute step count.

[e]PMSC: per minute step count.

[f]SC: step count.

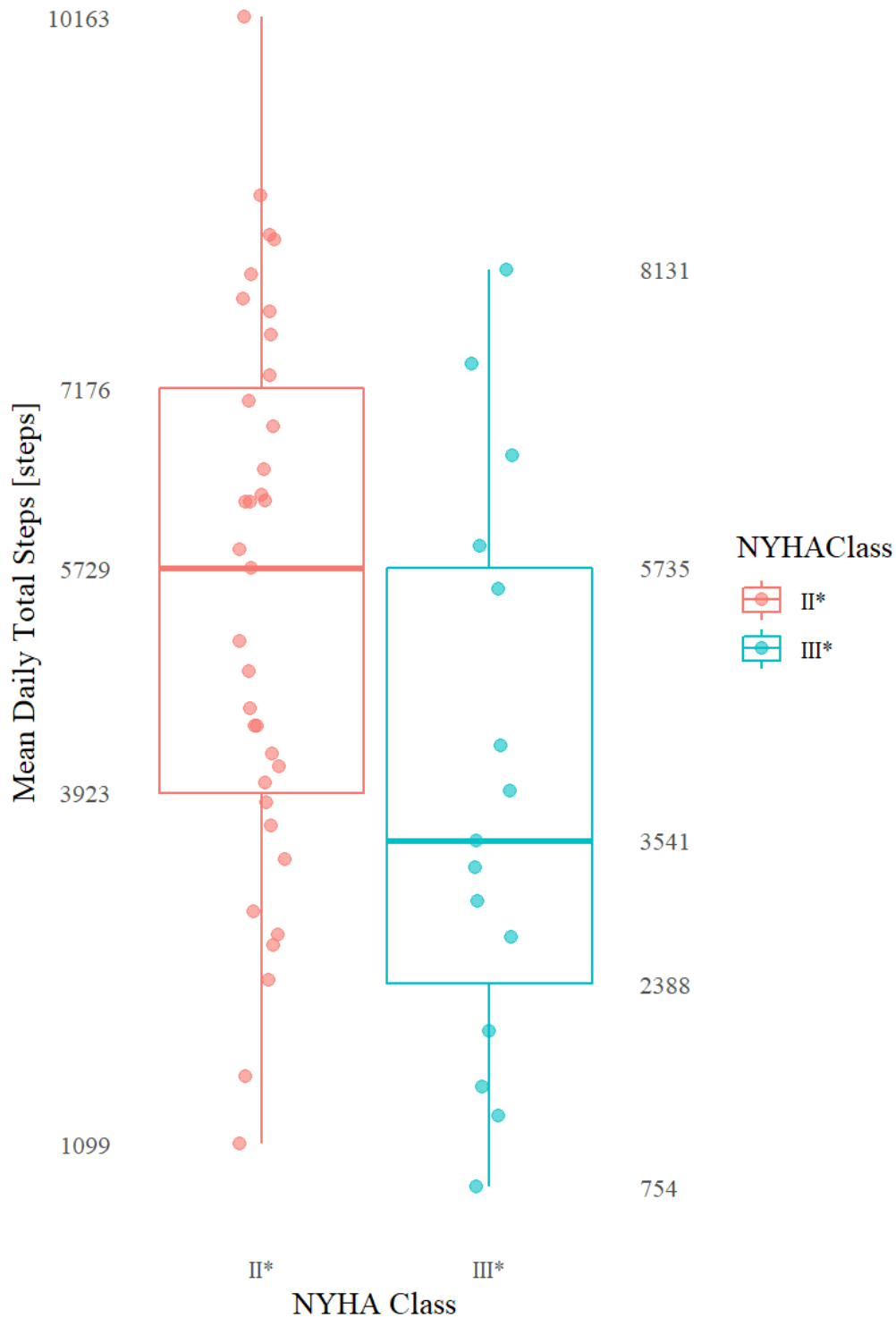**Figure 3.** Boxplots (min, Q1, median, Q3, max) of mean daily total steps for each NYHA class group.

Figure 4. Boxplots (min, Q1, median, Q3, max) of mean daily per minute step count maximums for each NYHA class group.
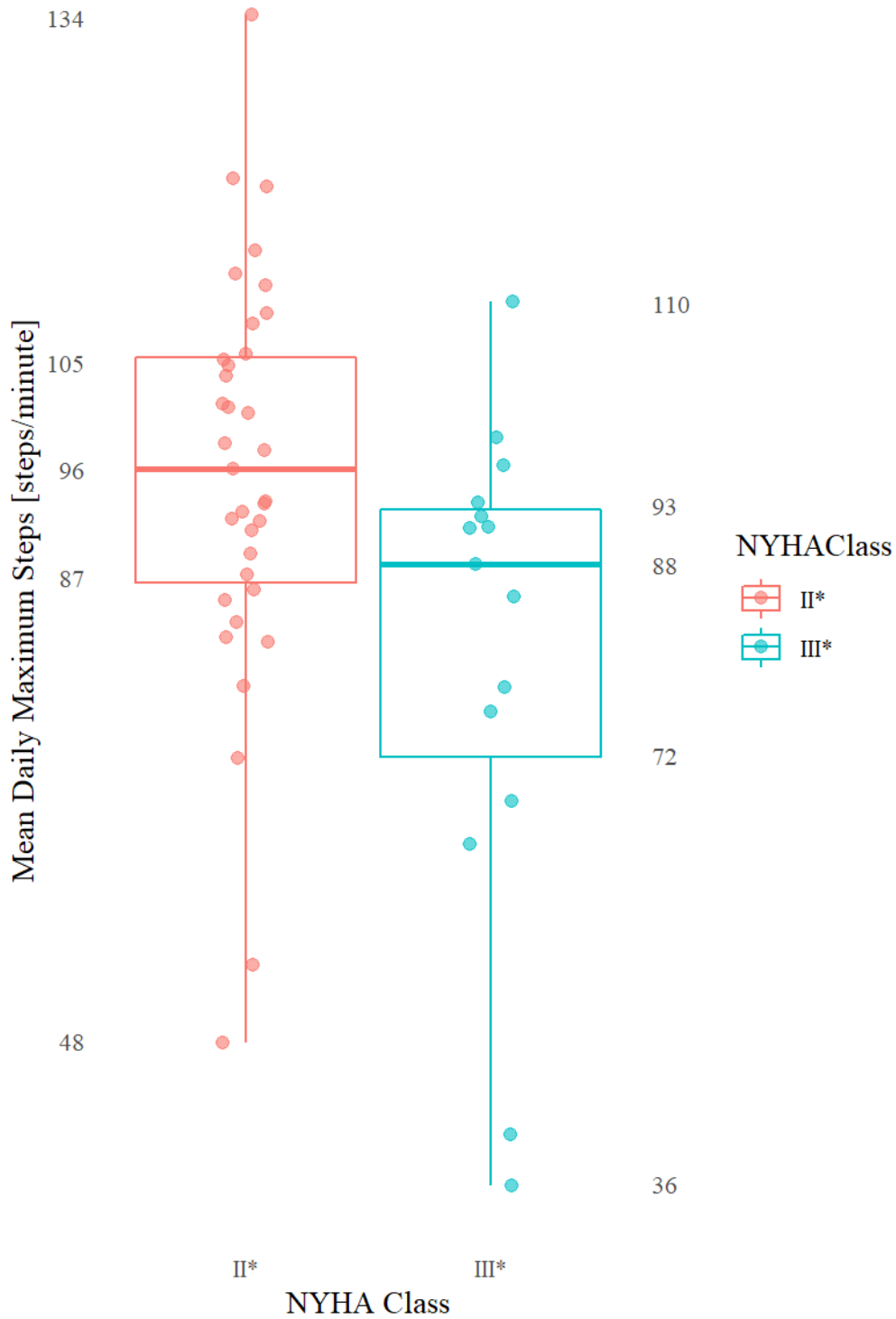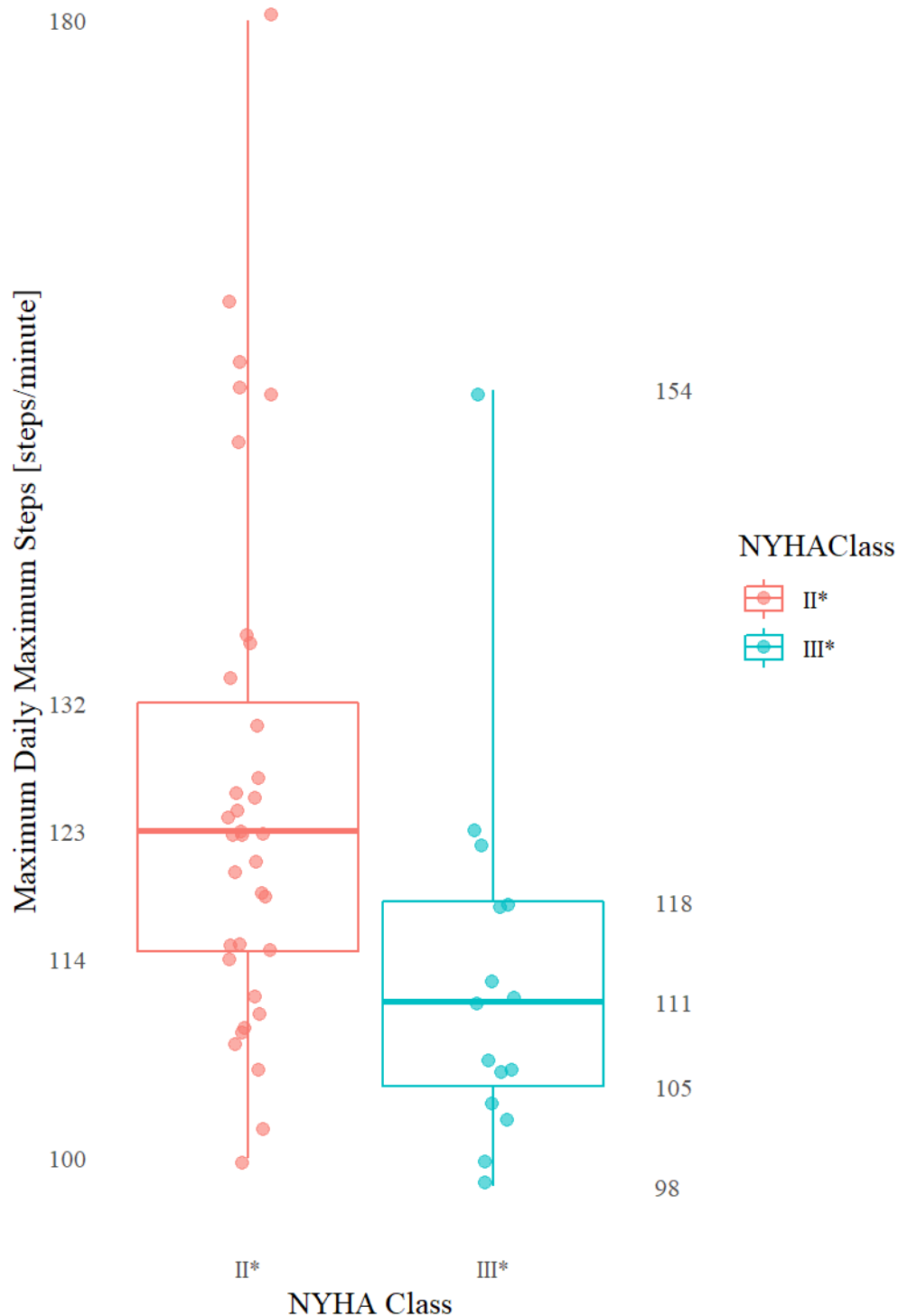
Figure 5. Boxplots (min, Q1, median, Q3, max) of maximum daily per minute step count maximums for each NYHA class group.

Discussion

Principal Findings

This study, using an independent, larger group of participants, replicated and validated the findings of our previous pilot study: that the daily free-living step counts of patients with HF exhibiting NYHA class II versus class III symptoms (ie, group II* vs group III*) are statistically different [18].

Specifically, HF patients categorized as NYHA II* and III* differed significantly (at the 5% level of significance) in their mean of daily total step counts (group medians: 5729 vs 3541; P=.04), maximum of daily total step counts (10792 vs 5904; P=.03), mean of daily mean step counts (4.0 vs 2.5; P=.04), as well as by their mean (96.1 vs 88.1; P=.02) of daily per minute step count maximums. These same patients differed significantly (at the .01% level of significance) by their maximum of daily per minute step count maximums (123.0 vs 111.0; P=.004,

respectively). The distribution of the per minute step counts by NYHA class—including only all nonzero per minute step count values—is shown in Figure 1. The daily step count results mimicked the 2-week overall step count values.

A total of 10,000 (steps/day) is often recommended as the daily step target for healthy adults, although in practice "many people can only achieve about slightly more than half of the daily step goal" with a meta-analysis of studies revealing ranges between 5300 and 6700 daily steps [29]. Persons who average <5000 (steps/day) are considered to be living a sedentary lifestyle, with persons averaging between 5000 to 7499 (steps/day) living a "low active" lifestyle [30,31]. Ayabe et al, based on a study of 77 cardiac rehab patients aged 46 to 88 years, recommended daily step targets of 6500 to 8500 (steps/day) for the secondary prevention of cardiovascular disease [32]. The NYHA group II* patients in our study, whose group median was a grand mean of 5729 (steps/day), achieved what would be considered a "low active" lifestyle near the bottom of the average daily step range of healthy adults and below the prevention target. In contrast, the NYHA group III* patients in our study, with a grand mean of 3541 (steps/day) (group median), fell well within the "sedentary" lifestyle range, well below the expected average daily step range of healthy adults and well below the secondary prevention target. Furthermore, at their peak within the 2-week study period—indicated by the maximum daily per minute step count total of 5904 (steps/day; group median)—the NYHA group III* patients never exceeded the "low active" lifestyle range neither did they come near to achieving the secondary prevention target, let alone the 10,000 (steps/day) target. In fact, at their peak, over the 2 weeks, the NYHA group III*'s maximum daily step count (group median: 5904 [steps/day]) only barely exceeded group II*'s grand mean step count (group median: 5729 [steps/day]). The NYHA group II* in comparison achieved a maximum daily per minute step count total of 10,792 (steps/day; group median): above both the secondary prevention target and the 10,000 (steps/day) target. Taken together, these numbers appear to quantitatively demonstrate a "marked limitation of physical activity" for patients with NYHA class III compared with a more "slight limitation of physical activity" for patients with NYHA class II (both corresponding to their respective NYHA functional classification criteria [17]).

As for the general shape of the step count distributions of the NYHA group II* versus III* patients, visual inspection of Figure 1 strongly suggests that there is a difference in the activity patterns of patients, for example, a longer, fatter tail for the group II* patients. Quantitatively, however, we failed to extract many meaningful insights into the shape of the activity distribution. The 1st, 2nd, and 3rd quartile (and by extension IQR), for example, were all found to be fairly consistently 0 for all patients, that is, 0's typically accounted for more than 75% (1,080/1,440) of the data points for any given patient day. This is because, unfortunately, the activity tracker used in this study records 0's both when a patient is not active and when the patient is simply not wearing the tracker. Not only does this make it difficult to ascertain if a 0-step count indicates lack of activity or patient's lack of adherence but it also means that we are unable to remove the excess 0's introduced into the apparent distribution as a result of a participant's lack of adherence to the tracker.

In light of the challenge introduced by the tracker selection, it is curious that in comparing the step count intensity measures (ie, maximum and mean daily aggregated per minute step count), the maximum daily per minute step count maximum values for each patient group was found to be notably more statistically significant compared with the other intensity measures. Of course, because of the nature of the statistic, metrics involving maximums would naturally be least susceptible to the ambiguous 0 per minute step count values. We suggest that this may be contributing to the daily maximum values appearing as more strongly differentiating between the 2 NYHA groups.

There are, however, other explanations for the phenomenon detailed above, including differences in accuracy of activity trackers at different step-intensity levels. Activity trackers, including the Fitbit, have been shown to be sufficiently accurate for research purposes [33-39]; however, several researchers have reported a degradation of accuracy in these devices (including the Fitbit Flex used in this study) at low and medium step cadences [35,37,39]. For example, An et al found that the accuracy of the device used in our study varied between 6.2% and 11.4% at the low and medium treadmill speeds (2 to 4 mph) that they tested but improved to 4% at the highest speed tested (5 mph) [37]. It is possible, therefore, that the more accurate recordings at high intensity levels simply makes it more possible to differentiate between the step counts of patients in each group regardless of the effect of superfluous 0-step count values.

Alternatively, it is also reasonable that the overall step count maximum, by capturing a patient's peak exercise capacity, might produce a more reliable (detectable) measure of the "limitation of physical activity" experienced by a patient in daily life and thus help differentiate more consistently between NYHA classes (compared with a simple mean or sum of a patient's activity over a said day). For example, previous in-laboratory studies observing patients performing a 6-minute walk test have been found that, on average, patients with the relatively higher NYHA class II spend more time (56%) at higher step intensities (>120 steps/minute) compared with patients with NYHA class III (24% of overall time) and vice-versa at lower step intensities (12% vs 36% of overall time at <100 steps/minute) [34]. It might just be that peak exercise generally may simply be a more consistent way of gaining insight into a patient's NYHA class than their average activity level.

Strengths and Limitations

A major limitation of this study is the grouping methodology used to reclassify patients who are assigned a borderline/mixed NYHA class to make them fit within the traditional 4-class NYHA classification system. The approach we used, although logically reasonable, has no demonstrated scientific support. Furthermore, the data being sourced as a convenience sample at the same single site, that is, consecutively recruited from the TGH Heart Function, represent a limitation of this study with regard to our objective of generalizing our findings. Our analysis was also limited as it did not include any patients with NYHA class I or IV. Although these are not typically as difficult to classify as NYHA class II or III patients, analysis of all 4 NYHA

classes would have potentially provided additional useful insight into the true underlying relationship between step count and NYHA class. Knowing exactly how step count and NYHA class are related may be tremendously valuable if it allows us to assess or predict NYHA class or gradation changes in NYHA class for a patient using their step count. We suggest that this might be the subject of an important future study.

The most significant limitation of our study, however, was the step tracker utilized, as it introduced significant ambiguity into the 0 per minute step count values which comprised most of each patient's step data stream. Zero values accounted for a mean of 87.3% (SD 4.9%) of the 2-week data stream for each patient—accounting for as much as 97.7% of the 2-week data stream for one of the patients. In fact, when looking at the 2-week period as a whole, they accounted for at least 76.7% of all the data points for any given patient. The complete breakdown is shown in Figure 6. Unfortunately, the meaning of these 0 per minute step count values is ambiguous as the trackers used in this study record a 0 value not only during patient inactivity—for example, when a patient is sitting, sleeping, or generally not moving—but also when the patient was simply not wearing the device—for example, to charge it. As a result, it is challenging to accurately determine if a given series of zeroes indicates a pattern of low physical activity—presumably explanatory of an NYHA class—or simply a pattern of no device usage—essentially introducing noise into the physical activity signal. This limits our ability to precisely quantify the distribution of the activity/inactivity of patients, especially as it is as of yet unclear how much importance patient inactivity (vs patient activity) should be accorded when it comes to capturing "physical activity limitation" and by extension the NYHA functional class. Investigations into how to disambiguate between inactive versus disengaged/nonadherence time in pedometer-like trackers would be hugely beneficial to help researchers correct for the effect of nonadherent time in the captured free-living step data distribution, especially if we are to better understand the actual true relationship between free-living activity and NYHA functional classification. At the very least, we recommend that future researchers strongly consider using an activity tracker that clearly disambiguates between inactivity and patient disengagement or provides an additional data stream that would support some reliable objective means of performing the disambiguation.

Figure 6. Number of zero step count minutes as a percentage of the total two-week data stream for each patient.

```
The decimal point is at the |
i.e. 76 | 7 represents 76.7%

76 | 78
78 | 9
80 | 2728
82 | 13678
84 | 022605688
86 | 03902226
88 | 024846
90 | 164668
92 | 14056
94 | 9
96 | 027
```

Conclusions

On average, patients exhibiting NYHA II versus NYHA III symptoms are expected to exhibit "low active" versus "sedentary" lifestyles with (1) mean daily step count totals around 5729 (steps/day) versus 3541 (steps/day; group medians)—in the case of patients exhibiting NYHA III symptoms less than the 5300 to 6500 (steps/day) expected of typical healthy adults and in the case of patients exhibiting NYHA II symptoms only barely within the same range and (2) maximum daily step count totals of 5904 versus 10,792 (steps; II vs III group medians)—compared with the healthy target of 10,000+ average (steps/day). These findings validate our previous pilot study and point to limitations in daily physical activity beyond those found in normal healthy adults. In addition, consistent with laboratory tests, patients exhibiting NYHA class III symptoms are on average expected to exhibit lower step count intensities during free-living with (group medians II vs III) (1) mean (2-week) daily mean step counts of 4.0 versus 2.5 (steps/minute), (2) a mean daily per minute step count maximums of 88.1 versus 96.1 (steps/minute), and (3) a maximum daily per minute step count maximums of 111.0 versus 123.0 (steps/minute).

The discovery of additional significant aggregate measures raises several questions, among them are the following: What is the exact underlying relationship between NYHA functional class and step count? What features of the step count waveform are most associated or correlated with NYHA functional class? These questions will no doubt feature as the subjects of future studies, but the findings of this study are an important milestone on the road to an objective means of assessing HF functional classification on our continuing quest to improve outcomes of patients with the burdensome and costly disease, that is, congestive HF.

Acknowledgments

Conflicts of Interest

SB owns shares of the manufacturer of the experimental device used in this study (Fitbit). The other coauthors have no conflicts of interest to disclose.

References

1. Savarese G, Lund L. Global public health burden of heart failure. Card Fail Rev 2017 Apr;3(1):7-11 [FREE Full text] [doi: 10.15420/cfr.2016:25:2] [Medline: 28785469]

2. Mehra M, Butler J. Heart failure: a global pandemic and not just a disease of the west. Heart Fail Clin 2015 Oct;11(4):xiii-xiiv. [doi: 10.1016/j.hfc.2015.08.001] [Medline: 26462110]

3. Heart and Stroke Foundation. Heart and Stroke Foundation. 2016. 2016 Report on the Health of Canadians: The Burden of Heart Failure URL: https://www.heartandstroke.ca/-/media/pdf-files/canada/2017-heart-month/heartandstroke-reportonhealth-2016.ashx?la=en&hash=0478377DB7CF08A281E0D94B22BED6CD093C76DB [accessed 2016-10-29] [WebCite Cache ID 706UliccA]

4. Lawrence S. Marketwired. Canada is failing our heart failure patients - Heart and Stroke Foundation of Canada [WebCite Cache ID 706U7G8oI]

5. Ezekowitz J, O'Meara E, McDonald M, Abrams H, Chan M, Ducharme A, et al. 2017 Comprehensive update of the Canadian Cardiovascular Society guidelines for the management of heart failure. Can J Cardiol 2017 Dec;33(11):1342-1433. [doi: 10.1016/j.cjca.2017.08.022] [Medline: 29111106]

6. Ziaeian B, Fonarow G. Epidemiology and aetiology of heart failure. Nat Rev Cardiol 2016 Dec;13(6):368-378 [FREE Full text] [doi: 10.1038/nrcardio.2016.25] [Medline: 26935038]

7. Kristensen S, Køber L, Jhund P, Solomon S, Kjekshus J, McKelvie R, et al. International geographic variation in event rates in trials of heart failure with preserved and reduced ejection fraction. Circulation 2015 Jan 6;131(1):43-53. [doi: 10.1161/CIRCULATIONAHA.114.012284] [Medline: 25406306]

8. Velagaleti R, Vasan R. Heart failure in the twenty-first century: is it a coronary artery disease or hypertension problem? Cardiol Clin 2007 Nov;25(4):487-95; v [FREE Full text] [Medline: 18063154]

9. Lee D, Gona P, Vasan R, Larson M, Benjamin E, Wang T, et al. Relation of disease pathogenesis and risk factors to heart failure with preserved or reduced ejection fraction: insights from the framingham heart study of the national heart, lung, and blood institute. Circulation 2009 Jun 23;119(24):3070-3077 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.108.815944] [Medline: 19506115]

10. Public Health Agency of Canada. Ottawa; 2018 May. Report from the Canadian Chronic Disease Surveillance System: Heart Disease in Canada, 2018 URL: https://www.canada.ca/content/dam/phac-aspc/documents/services/publications/diseases-conditions/report-heart-disease-canada-2018/pub1-eng.pdf [WebCite Cache ID 75UuX4oGO]

11. Public Health Agency of Canada. Public Health Infobase. Ottawa Canadian Chronic Disease Surveillance System Internet URL: https://infobase.phac-aspc.gc.ca/CCDSS-SCSMC/data-tool/?l=eng&HRs=00&DDLV=11&DDLM=PREV&CBVS=on&Age=40andOver&1=M&2=F&DDLFrm=2000&DDLTo=2015&=10&VIEW=2 [accessed 2019-02-08]

12. Raphael C, Briscoe C, Davies J, Whinnett ZI, Manisty C, Sutton R, et al. Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. Heart 2007 Apr;93(4):476-482 [FREE Full text] [doi: 10.1136/hrt.2006.089656] [Medline: 17005715]

13. Stein K, Mittal S, Merkel S, Meyer TE. Baseline physical activity and NYHA classification affects future ventricular event rates in a Ggeneral ICD population. J Card Fail 2006 Aug;12(6):S58. [doi: 10.1016/j.cardfail.2006.06.203]

14. Ahmed A, Aronow W, Fleg J. Higher New York Heart Association classes and increased mortality and hospitalization in patients with heart failure and preserved left ventricular function. Am Heart J 2006 Feb;151(2):444-450 [FREE Full text] [doi: 10.1016/j.ahj.2005.03.066] [Medline: 16442912]

15. National Heart Foundation of Australia and the Cardiac Society of Australia and New Zealand (Chronic Heart Failure Guidelines Expert Writing Panel). 2011 Oct. New York Heart Association (NYHA) classification URL: http://www.heartonline.org.au/media/DRL/New_York_Heart_Association_(NYHA)_classification.pdf [accessed 2019-05-06] [WebCite Cache ID 78A4Q0vu1]

16. Bennett JA, Riegel B, Bittner V, Nichols J. Validity and reliability of the NYHA classes for measuring research outcomes in patients with cardiac disease. Heart Lung 2002;31(4):262-270. [Medline: 12122390]

17. American Heart Association. 2015. Classes of Heart Failure URL: http://www.heart.org/HEARTORG/Conditions/HeartFailure/AboutHeartFailure/Classes-of-Heart-Failure_UCM_306328_Article.jsp#.WvyuQYgvyiN [accessed 2016-10-30] [WebCite Cache ID 6zT3C5Rpx]

XSL·FO
RenderX

18. Moayedi Y, Abdulmajeed R, Duero Posada J, Foroutan F, Alba A, Cafazzo J, et al. Assessing the use of wrist-worn devices in patients with heart failure: feasibility study. JMIR Cardio 2017 Dec 19;1(2):e8. [doi: 10.2196/cardio.8301]

19. Fitbit Inc. 2016. Fitbit Official Site for Activity Trackers & More URL: https://www.fitbit.com/en-ca/home [accessed 2018-05-17] [WebCite Cache ID 6zTITrK95]

20. Bromberg S. GitHub. 2015. Googlefitbit URL: https://github.com/simonbromberg/googlefitbit [accessed 2019-04-29] [WebCite Cache ID 782HJo1P3]

21. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2017. URL: https://www.r-project.org [accessed 2019-05-01] [WebCite Cache ID 782HRiPHK]

22. RStudio Team. RStudio: Integrated Development Environment for R. Boston, MA; 2015. URL: https://www.rstudio.com/ [accessed 2019-05-01] [WebCite Cache ID 782HcI2y3]

23. Wickham H. A layered grammar of graphics. J Comput Graph Stat 2010 Jan;19(1):3-28. [doi: 10.1198/jcgs.2009.07098]

24. Arnold JB. cran.r-project. 2017. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2' URL: https://cran-r-project.org/package=ggthemes [accessed 2019-04-30] [WebCite Cache ID 782Hpuxu6]

25. Wickham H. The split-apply-combine strategy for data analysis. J Stat Soft 2011;40(1):- [FREE Full text] [doi: 10.18637/jss.v040.i01]

26. Wickham H, Francois R, Henry L, Müller K. cran.r-project. 2017. dplyr: A Grammar of Data Manipulation URL: https://cran-r-project.org/package=dplyr [accessed 2019-04-30] [WebCite Cache ID 782Hv0Idp]

27. Wickham H. Reshaping data with the reshape package. J Stat Soft 2007;21(12):- [FREE Full text] [doi: 10.18637/jss.v021.i12]

28. Hester J. cran.r-project. 2017. glue: Interpreted String Literals URL: https://cran-r-project.org/web/packages/glue/index.html [accessed 2019-04-30] [WebCite Cache ID 782I9aGci]

29. Choi BC, Pak AW, Choi JC, Choi EC. Daily step goal of 10,000 steps: a literature review. Clin Invest Med 2007;30(3):E146-E151. [Medline: 17716553]

30. Tudor-Locke C, Hatano Y, Pangrazi RP, Kang M. Revisiting. Med Sci Sports Exerc 2008 Jul;40(7 Suppl):S537-S543. [doi: 10.1249/MSS.0b013e31817c7133] [Medline: 18562971]

31. Tudor-Locke C, Craig C, Thyfault J, Spence J. A step-defined sedentary lifestyle index: <5000 steps/day. Appl Physiol Nutr Metab 2013 Feb;38(2):100-114. [doi: 10.1139/apnm-2012-0235] [Medline: 23438219]

32. Ayabe M, Brubaker P, Dobrosielski D, Miller H, Kiyonaga A, Shindo M, et al. Target step count for the secondary prevention of cardiovascular disease. Circ J 2008;72(2):299-303 [FREE Full text] [doi: 10.1253/circj.72.299]

33. Storm FA, Heller BW, Mazzà C. Step detection and activity recognition accuracy of seven physical activity monitors. PLoS One 2015;10(3):e0118723 [FREE Full text] [doi: 10.1371/journal.pone.0118723] [Medline: 25789630]

34. Jehn M, Prescher S, Koehler K, von Haehling S, Winkler S, Deckwart O, et al. Tele-accelerometry as a novel technique for assessing functional status in patients with heart failure: feasibility, reliability and patient safety. Int J Cardiol 2013 Oct 12;168(5):4723-4728. [doi: 10.1016/j.ijcard.2013.07.171] [Medline: 23962782]

35. Evenson K, Goto M, Furberg R. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]

36. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]

37. An H, Jones G, Kang S, Welk G, Lee JM. How valid are wearable physical activity trackers for measuring steps? Eur J Sport Sci 2017 Apr;17(3):360-368. [doi: 10.1080/17461391.2016.1255261] [Medline: 27912681]

38. El-Amrawy F, Nounou MI. Are currently available wearable devices for activity tracking and heart rate monitoring accurate, precise, and medically beneficial? Healthc Inform Res 2015 Oct;21(4):315-320 [FREE Full text] [doi: 10.4258/hir.2015.21.4.315] [Medline: 26618039]

39. Thorup C, Andreasen J, Sørensen EE, Grønkjær M, Dinesen B, Hansen J. Accuracy of a step counter during treadmill and daily life walking by healthy adults and patients with cardiac disease. BMJ Open 2017 Dec 31;7(3):e011742 [FREE Full text] [doi: 10.1136/bmjopen-2016-011742] [Medline: 28363918]

Abbreviations

CHD: coronary heart disease
EF: ejection fraction
HF: heart failure
HFpEF: heart failure with preserved ejection fraction
HFrEF: heart failure with reduced ejection fraction
HT: hypertension
IQR: interquartile range
LVEF: left ventricular ejection fraction
NYHA: New York Heart Association
TGH: Toronto General Hospital

XSL·FO
RenderX

XSL•FO
RenderX

<u>Original Paper</u>

Provider- and Patient-Related Barriers to and Facilitators of Digital Health Technology Adoption for Hypertension Management: Scoping Review

Ramya Sita Palacholla[1,2,3], MD, MPH; Nils Fischer[1,2], MPH; Amanda Coleman[4], MPH; Stephen Agboola[1,2,3], MD, MPH; Katherine Kirley[4], MD, MS; Jennifer Felsted[1,2], PhD; Chelsea Katz[4], MBA; Stacy Lloyd[4], MPH; Kamal Jethwani[1,2,3], MD, MPH

[1]Partners HealthCare Pivot Labs, Boston, MA, United States

[2]Harvard Medical School, Boston, MA, United States

[3]Massachusetts General Hospital, Boston, MA, United States

[4]American Medical Association, Chicago, IL, United States

Corresponding Author:
Ramya Sita Palacholla, MD, MPH
Massachusetts General Hospital
55 Fruit St
Boston, MA, 02114
United States
Phone: 1 8579199889
Email: rpalacholla@mgh.harvard.edu

Abstract

Background: The uptake of digital health technology (DHT) has been surprisingly low in clinical practice. Despite showing great promise to improve patient outcomes and disease management, there is limited information on the factors that contribute to the limited adoption of DHT, particularly for hypertension management.

Objective: This scoping review provides a comprehensive summary of barriers to and facilitators of DHT adoption for hypertension management reported in the published literature with a focus on provider- and patient-related barriers and facilitators.

Methods: This review followed the methodological framework developed by Arskey and O'Malley. Systematic literature searches were conducted on PubMed or Medical Literature Analysis and Retrieval System Online, Cumulative Index to Nursing and Allied Health Literature, and Excerpta Medica database. Articles that reported on barriers to and/or facilitators of digital health adoption for hypertension management published in English between 2008 and 2017 were eligible. Studies not reporting on barriers or facilitators to DHT adoption for management of hypertension were excluded. A total of 2299 articles were identified based on the above criteria after removing duplicates, and they were assessed for eligibility. Of these, 2165 references did not meet the inclusion criteria. After assessing 134 studies in full text, 98 studies were excluded (full texts were either unavailable or studies did not fulfill the inclusion criteria), resulting in a final set of 32 articles. In addition, 4 handpicked articles were also included in the review, making it a total of 36 studies.

Results: A total of 36 studies were selected for data extraction after abstract and full-text screening by 2 independent reviewers. All conflicts were resolved by a third reviewer. Thematic analysis was conducted to identify major themes pertaining to barriers and facilitators of DHT from both provider and patient perspectives. The key facilitators of DHT adoption by physicians that were identified include ease of integration with clinical workflow, improvement in patient outcomes, and technology usability and technical support. Technology usability and timely technical support improved self-management and patient experience, and positive impact on patient-provider communication were most frequently reported facilitators for patients. Barriers to use of DHTs reported by physicians include lack of integration with clinical workflow, lack of validation of technology, and lack of technology usability and technical support. Finally, lack of technology usability and technical support, interference with patient-provider relationship, and lack of validation of technology were the most commonly reported barriers by patients.

Conclusions: Findings suggest the settings and context in which DHTs are implemented and individuals involved in implementation influence adoption. Finally, to fully realize the potential of digitally enabled hypertension management, there is

a greater need to validate these technologies to provide patients and providers with reliable and accurate information on both clinical outcomes and cost effectiveness.

Introduction

Digital health technologies (DHTs) have the potential to support active self-management of chronic conditions via education, monitoring and support, timely feedback, and remote access to health professionals [1]. When designed and implemented successfully, digital health interventions offer an opportunity to support the quadruple aim of health care by improving health outcomes, increasing patient experience, reducing health care costs, and improving clinician satisfaction [2]. The American Medical Association (AMA) defines digital health tools as those systems and solutions that engage patients for clinical purposes, collect, organize, interpret, use clinical data, and manage outcomes and other measures of care quality including telemedicine and telehealth, mobile health, wearables, remote monitoring, and apps [3]. The AMA digital health survey classifies digital health solutions into 7 categories: remote monitoring for efficiency, remote monitoring and management for improved care, clinical decision support, patient engagement, televisits, point-of-care, and tools providing consumer access to clinical data [3].

One-third of the US population has hypertension (85.7 million adults) [4] and the economic burden is close to US $ 53 billion dollars annually [5]. Despite having access to effective drugs for lowering blood pressure (BP), BP control in a vast majority of patients remains suboptimal [5], owing to infrequent monitoring of BP [6], low medication adherence by patients [7], and clinical inertia [8]. DHTs for hypertension management, such as telemonitoring programs, enhance self-monitoring as they allow for BP readings and clinical information to be shared with health care professionals in real time [9]. Remote monitoring for hypertension has been shown to improve medication adherence [10], optimize BP control [11], and reduce use of health care resources [12].

Although the shift to a value-based care system has encouraged the adoption and use of DHT to manage hypertension, the uptake of DHTs has been surprisingly low in clinical practice [13]. In addition, to our knowledge, there is limited information on the factors that influence adoption of digital health from the perspectives of both patients and providers. Previously published literature includes surveys of providers that cite factors influencing DHT adoption such as organizational and financial barriers [14]. Previous systematic reviews of telemedicine for hypertension management report increased access to health services, improved health and quality outcomes, and enhanced patient knowledge and involvement in disease management as strong facilitators of DHT usage in health care settings [13,15]. This review provides a comprehensive summary of facilitators and barriers to adopting digital health for hypertension management with a specific focus on the perspectives of providers and patients.

Methods

Literature Search

This scoping review was conducted using the methodological framework developed by Arskey and O'Malley [16]. The Arksey and O'Malley framework is particularly suited to address broad research questions and can help map the current literature, extract key concepts and themes, and identify gaps. The Arksey and O'Malley framework has several steps including (1) identifying the broad research question, (2) study selection using inclusion or exclusion criteria on the basis of familiarity with the topic of interest, (3) sorting the extracted data from studies into themes and patterns, and (4) collating key themes and issues [16]. The primary research question guiding this review was the following: *What are the barriers and facilitators of digital health adoption for hypertension management?*

Structured literature searches were conducted using 3 databases to identify relevant studies from 2008 to 2017: PubMed or Medical Literature Analysis and Retrieval System Online, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Excerpta Medica database (EMBASE). Medical subject headings (MeSH) and selected keywords were searched using Boolean operator *OR* and these groups were combined using another Boolean operator *AND*. Keywords used include (1) *hypertension* (MeSH), *hypertensi*, (2) *mobile applications* (MeSH), *mobile device*, (3) *electronic health records* (MeSH), *personal electronic health record*, (4) *decision support systems, decision support*, (5) *remote monitoring* (MeSH), (6) *providers* (MeSH)*, clinician*. The detailed search strategies for PubMed have been provided as an example (see Multimedia Appendix 1). At first, 2 reviewers, with subject matter and methodological expertise, independently reviewed all abstracts identified by the searches and conflicts were resolved by a third reviewer. Then, 2 reviewers screened the full texts to select the final studies to be included in the review. Cohen kappa test revealed an agreement score of 0.75 between the reviewers. Per Landis and Koch, this agreement score could be categorized as *substantial agreement* between the reviewers [17].

All articles retrieved were screened using the following inclusion criteria: (1) reported on adoption barriers and/or facilitators of digital health solutions, as defined by the AMA, that were provider- or patient-related, (2) focused on hypertension management, (3) published in English, and (4) published between 2008 and 2017. Studies were excluded if they (1) did not report on barriers or facilitators of digital health, (2) described barriers or facilitators exclusively for nonclinical staff such as pharmacists, (3) were editorials or reviews for editorials,

epidemiological studies, and protocols, (4) provided insights on acute management of hypertension in perioperative or intensive care settings, or (5) if full texts were unavailable. The authors also conducted a gray literature search (including conference proceedings) through a Web search engine. In addition, 4 articles were handpicked on the basis of the same inclusion criteria used for articles selected via literature databases.
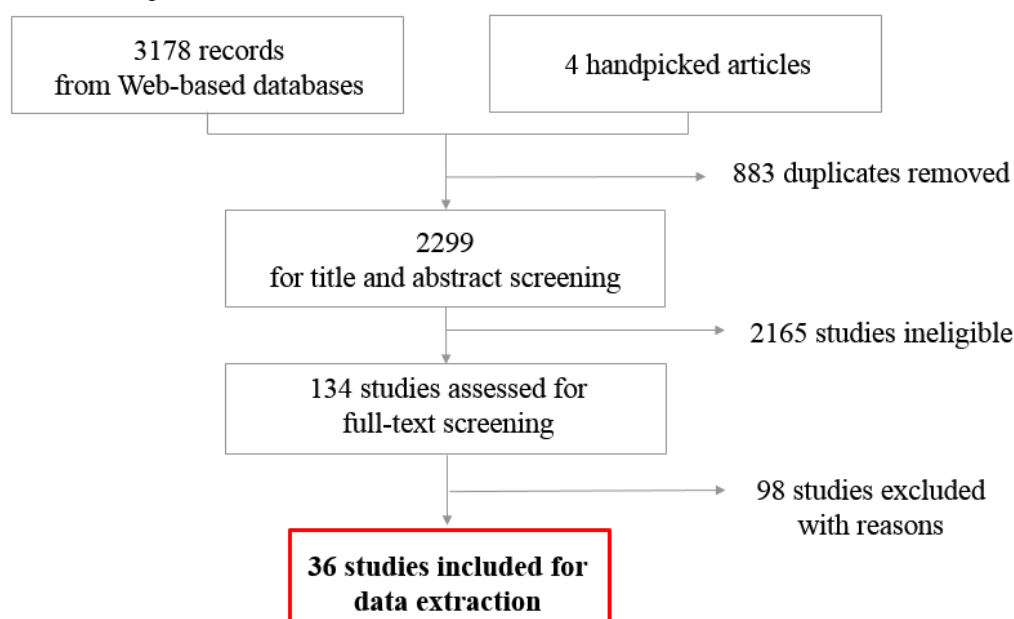
Thematic Analysis

The selected papers were reviewed to extract relevant data. A data extraction template developed by the authors was used to extract key information and concepts from the included studies and the template included the following constructs: the geography, study design, program setting, disease conditions (in addition to hypertension), study objectives, sample description, sample size, digital health category, design features of the intervention, clinical outcomes, cost outcomes, patient experience, provider experience, patient-related barriers and facilitators, and provider-related barriers and facilitators. Descriptive and inductive thematic analyses were conducted for identifying major themes pertaining to barriers and facilitators of DHT adoption. For the analysis of the text passages from the included articles, the inductive thematic analysis was conducted as described by Braun and Clarke [18]. We developed our own a priori framework to categorize barriers into the following 4 categories: (1) provider-related facilitators, (2) provider-related barriers, (3) patient-related barriers, (4) and patient-related facilitators. This analytic process involved reading and rereading of the selected papers, systematically identifying and naming the unit of meaning with codes (words or sets of words that provide a meaning label), and then searching for patterns in the data and organizing the data (smaller themes or codes) into larger themes representing the main ideas and their relationships. Themes were then reviewed by the team and representative data elements were selected to demonstrate the salient themes. At first, 2 investigators (RP and NF) independently performed the initial coding of the first transcript. This coding was then reviewed by the third reviewer (AC). The codes were then reviewed and discussed with the team including senior researchers in the field, providers, and other subject matter experts. Later, 2 reviewers (RP and NF) then recoded all papers, integrating feedback from the team into

the coding structure. A final codebook was created using Microsoft Office Excel (version 1808) on the basis of the consensus of the 3 investigators (RP, NF, and AC). During this process, any discrepancies in coding were discussed and resolved among all investigators. Furthermore, any questions about meaning and interpretation of themes were discussed among the team members and resolved through consensus.

Results

Overview

A total of 2299 titles and abstracts from PubMed, CINAHL, EMBASE, and 4 handpicked articles from the supplementary gray literature search were assessed for eligibility after removing duplicates (see Figure 1). Of these, 2165 references did not meet the inclusion criteria. After assessing 134 studies in full text, 98 studies were excluded (full texts were either unavailable or studies did not fulfill the inclusion criteria). A total of 36 studies satisfied the inclusion criteria, including the 4 handpicked articles. The articles included in this review were published between 2008 and 2017, with a majority (n=30) published after 2010. Studies were published across the following countries: United States (n=21), United Kingdom (n=4), Canada (n=3), Finland (n=1), Sweden (n=1), Italy (n=1), Taiwan (n=1), Malaysia (n=1), South Korea (n=1), Kenya (n=1), and Germany (n=1). DHTs included in this review were classified into categories as defined by the AMA: remote monitoring for efficiency (n=6), remote monitoring and management for improved care (n=19), clinical decision support (n=6), patient engagement (n=4), televisits or virtual visits (n=6), point-of-care(n=2), and tools providing consumer access to clinical data (n=1). Most studies were conducted in a primary care setting (n=30). A plurality of studies included qualitative assessments (n=15). Quantitative methodologies included randomized controlled trials (RCTs; n=14), nonrandomized trials (n=2), usability pilots (n=2), and pre and poststudies (n=2). In addition, 1 white paper was also included in this review. Multimedia Appendix 2 displays a summary of the studies included in this review. The results of the thematic analysis have been categorized as provider- and patient-related facilitators and barriers as detailed below. Tables 1 and 2 summarize all the themes.

Figure 1. Study selection flow diagram.

Table 1. Summary and frequency of provider-related themes and sub-themes identified from authors' thematic analysis of the 36 studies in this review. Most studies included in the review reported multiple themes. Frequency of a barrier or a facilitator=total number of occurrences of a facilitator or the barrier and total frequency of occurrences of facilitators and barriers.

Variable	Occurrences and frequency, n (%)
Facilitators[a]	
1. Ease of integration with clinical workflow [19-25]; Actionable data to provide timely interventions to patient [20,22,23,26]; Integration with clinical routine and less time-consuming tasks [20,21]; Care team support: opportunity for delegation and team-based care [19,20,24,25]	8 (33)
2. Improvement in patient health outcomes [20,23,25,27,28]; Technology prevalidated to improve outcomes [20,25,27]; Positive impact on patients and their self-management [20,27,28]; Better monitoring of patients to prevent negative outcomes [23]	5 (21)
3. Technology usability and technical support [29-36]; Technology requires minimal training [29,35,36]; Ease of use [29,30,35,36]; Adequate training support [31-33]	8 (33)
4. Financial factors [27,37]	2 (8)
5. Leadership and organizational support [38]	1 (4)
Barriers[b]	
1. Lack of integration with clinical workflow [9,19-21,24,25,30,39-41]; Lack of integration with electronic medical record [24]; Additional time-consuming tasks for providers [9,19-21,39-41]; Clinically irrelevant data [25,30]	10 (36)
2. Lack of validation of technology [14,32,38,42-45]; Concern over data accuracy [14,42-44]; Lack of evidence of improvement in patient outcomes [32,38,45]	7 (25)
3. Concern over data privacy and security [32]	1 (4)
4. Lack of technology usability and technical support [30,34,38-40,43,46]; Frequent technical issues [34,39] Lack of ease of use [30,39,40,43,46]; Long learning curve [38]	7 (25)
5. Lack of leadership and organizational support [32,40]	2 (7)
6. Increased patient anxiety [14]	1 (4)

[a]Total frequency of occurrences of facilitators=20.

[b]Total frequency of occurrences of barriers=28.

XSL•FO
RenderX

Table 2. Summary and frequency of patient-related themes and sub-themes identified from authors' thematic analysis of the 36 studies in this review. Most studies included in the review reported multiple themes. Frequency of barrier and facilitator=total number of occurrences of a facilitator or barrier and total frequency of occurrences of facilitators or barriers.

Variable	Occurrences and frequency, n (%)
Facilitators[a]	
1. Technology usability [19,24,30,34,36,46-50]; Ease of use [19,24,30,36,46,48-50]; Technical support [47,48]; Integration into patient's daily routine [46]	10 (29)
2. Positive impact on patient-provider communication [19,20,28,37,46,49-51]; Improved and more timely feedback from providers [19,20,28,37,46,49-51]; Shared decision making with providers [46]; Better preparation for clinic visits [28]	8 (24)
3. Improved self-management and patient experience [19,21,24,30,33,36,39,46,52]; Increased motivation to better manage health [36,39]; Increased access to health data [21,24,30,33,36,46]; Alleviation in anxiety from better monitoring of health data [19,52]	9 (26)
4. Reduction of in-office visits [19,21,24,25,37,41,52]	7 (20)
Barriers[b]	
1. Lack of technology usability and technical support [14,19,20,30,47,48]; Frequent technical glitches [14,19,21,50,53]; Lack of ease of use of system [43,50]; Patient not confident in using device [14,19,50]	9(41)
2. Interference with patient-provider relationship [19,20,37,42,47]; Fear of having less direct in-person communication with provider [19,37]; Lack of feedback from providers [42,47]; Disrupting feelings of independence [20,37]	5 (23)
3. Lack of validation of technology [19,43,47]	3 (14)
4. Increased patient anxiety [49,52]	2 (9)
5. Concern over data privacy and security [48]	1 (5)
6. Cost of digital health equipment [42,47]	2 (9)

[a]Total frequency of occurrences of facilitators=34.

[b]Total number of occurrences of barriers=19.

Facilitators of Digital Health Adoption

Provider Factors

Ease of Integration With Clinical Workflow

The findings suggest that integration of a new technology into the existing workflow of a provider strongly influences DHT adoption (n=2) [20,21]. Providers cited that having a care team to support DHT implementation as part of the clinical workflow was an important facilitator of adoption (n=4) [19,20,24,25]. Some studies found that providers were able to successfully adopt DHTs when the data that the DHT provided were actionable and could be readily utilized within preexisting clinical workflows to enable timeline intervention to improve patient outcomes (n=4) [20,22,23,26]. Providers were also attracted to DHTs that provided automatic alerts identifying the need for a change in medications or dosage [23], as they helped perform routine tasks faster (n=1).

Improvement in Patient Health Outcomes

Providers' beliefs regarding whether the technology improved clinical outcomes or engaged patients in self-management were among the most important considerations (n=3) [20,27,28] for embracing DHTs. In some instances, the DHTs that were validated in pilot and RCTs and shown to improve outcomes were perceived to be more acceptable to providers (n=4) [20,25,27,28]. Furthermore, providers valued their patients becoming more active and engaged in their own health (n=2) [20,28]. Finally, DHTs that enabled a more timely response to

elevated BP levels helped providers prevent adverse health outcomes in their patients by addressing the changes in BP levels in a timely manner (n=1) [23].

Technology Usability and Technical Support

Some studies reported that providers valued the simplicity and ease of use of a system (n=4) [29,30,35,36]. Furthermore, providers preferred DHTs that required minimal training (n=3) [29,35,36]. Providers valued adequate technical support when using DHTs as a part of their clinical workflow (n=3) [30,34,35].

Financial Factors

A few studies reported that financial incentives such as physician reimbursement for using DHTs in their clinical practice and cost savings as a result of implementing DHTs were important influencers of provider adoption (n=2) [27,37].

Leadership and Organizational Support

An organizational culture of innovation coupled with the presence of *physician champions* was cited as a factor influencing the adoption of DHTs in clinical settings, as it was often difficult for clinicians to implement DHTs without the support of their organization and leadership, particularly in terms of required budget and personnel (n=1) [38].

Patient Factors

Technology Usability and Technical Support

DHTs that were easy to use and included timely technical support [19,24,30,34,36,46,47,49,50] fostered patient

engagement (n=9). Older patients and those with less experience using technology reported that technical support was a facilitator (n=2) [47,48]. Patients valued solutions that were easy to integrate into their daily routines (n=1) [46]. Interventions were more easily adopted when they were culturally tailored for specific target populations (n=1) [34].

Improved Patient-Provider Communication

Improved communication with providers was a facilitator of adoption for patients. Some patients reported that DHTs enabled direct contact with their providers to share their health data and receive feedback [19,20,28,37,46,49-51]. Data sharing via DHTs helped patients better understand their care plans and promoted shared decision making [46]. DHTs improved visit preparation and accuracy of patient-provided information [28].

Improved Self-Management and Patient Experience

Patients were more likely to adopt DHTs that increased their motivation to manage their own conditions (n=2) [36,39]. Patients reported that being able to access and view their health data from their own device encouraged them to be more proactive about their health (n=6) [21,24,30,33,36,46]. Several studies reported greater patient satisfaction using DHTs for hypertension management (n=6) [19,36,37,47,48,52]. Some patients found that using DHTs to monitor their BP readings helped alleviate health-related anxiety (n=2) [19,52].

Reduction of Office Visits

The opportunity for patients to potentially avoid having to travel to the physician's office was reported as a facilitator of DHT adoption by patients in some studies (n=7) [19,21,24,25,37,41,52].

Barriers for Digital Health Adoption

Provider Factors

Lack of Integration With Clinical Workflow

Several studies reported the lack of integration of technology with clinical workflow as a major barrier to DHT adoption (n=6) [21,24,25,39-41]. The lack of care team resources available to successfully implement DHTs and perform additional tasks was highlighted by multiple studies (n=3) [19,20,24]. Too many additional tasks associated with implementing DHTs were reported to be problematic for several providers (n=1) [9].

Lack of Validation of Technology

Some providers cited concerns over accuracy of data as a potential road block to using home BP monitors on a wider scale (n=4) [14,42-44]. Another barrier to provider adoption was the lack of evidence or proof that DHTs improved patient outcomes (n=3) [32,38,45].

Concern Over Data Privacy and Security

One study reported that the lack of assurance of patient data security was a big concern for providers as well (n=1) [32].

Lack of Technology Usability and Technical Support

Another barrier frequently highlighted in the literature was the complexity of technologies (n=5) [30,39,40,43,46]. Frequent technical issues coupled with inadequate onsite support to resolve them were cited as reasons for discontinuing engagement

with DHTs (n=2) [34,39]. Furthermore, the learning curve associated with new DHTs made it difficult for providers to balance the use of a new system and keep up with their daily clinical routine (n=1) [38].

Lack of Organizational Support

Organizational factors, such as lack of leadership support for integrating technology in practice and budget constraints, delayed implementation of new DHTs (n=2) [32,40]. Hospital budgets were too constrained to gather additional resources necessary to implement DHTs as part of the clinical practice workflow (n=1) [32].

Increased Patient Anxiety

One study reported that providers were concerned that patients may be more anxious if they continuously monitored their BP data and believed excess data could be more harmful than useful for the patients (n=1) [14].

Patient Factors

Lack of Technology Usability and Technical Support

Technical issues such as password access, connectivity, and usability prevented patients from using DHTs (n=5) [14,19,20,30,48]. Patients often preferred DHTs that were easy to use regardless of technical skills and abilities and were less time consuming (n=2) [47,48]. Patients with impaired vision, low dexterity, and chronic conditions had difficulties adopting DHTs into their routine (n=3) [14,20,48].

Interference With Patient-Provider Relationship

Patients expressed concerns that using DHTs would interfere with their current in-person relationship with their providers (n=2) [19,37]. Another barrier that patients experienced was the lack of timely feedback from the provider when using DHTs with a provider-facing portal (n=2) [42,47]. In some cases, DHTs were viewed as an impediment to patients' feelings of independence as they were forced to share data with providers they may not want to (n=2) [19,20].

Increased Patient Anxiety

Some patients experienced anxiety from using DHTs (n=2) [49,52]. This anxiety stemmed from checking their BP too often and being unable to contact their provider directly and obtain timely feedback (n=2) [49,52].

Concern Over Data Privacy and Security

Patients were comfortable with access to health data being limited to only themselves and their providers. However, patients were concerned about the privacy of data shared via DHTs and were uncomfortable with the risk of a third party accessing their data [48].

Lack of Validation of Technology

In some studies, patients questioned the accuracy of the measurements and data recorded (BP readings) by DHTs [19,43,47].

Cost of Digital Health Equipment

The cost of digital health equipment was also cited as a barrier to adoption [42,47]. Some patients also expressed concern over being liable for cost of damage to the equipment [47].

Discussion

Principal Findings

This review contributes to existing literature by highlighting factors that enable or hinder the adoption of digital health solutions from the perspectives of both providers and patients. These results show that the key facilitators of DHT adoption by physicians include integration with clinical workflow 33% (8/24), ease of use 21% (5/24), improvement in patient outcomes 21% (5/24), financial factors 8% (2/20), and organizational support 4% (1/20). Technology usability and technical support 29% (10/35), positive impact on well-being and self-management 26% (9/35), improved patient-provider relationship 24% (8/35), and a reduction of in-office visits 20% (7/35) were most frequently reported facilitators for patients. The most frequently reported barriers to use of DHTs reported by physicians include lack of integration with workflow 36% (10/28), lack of validation of technology 25% (7/28), and lack of usability and support 25% (7/28). Finally, a lack of technology usability 41% (9/22), interference with the patient-provider relationship 23% (5/22), and lack of validation of technology 14% (3/22) were the top barriers reported by patients.

Although these findings highlight some common themes reported in previous work, there are several key differences and contributions from this study. A 2017 study by Mileski et al, examining the facilitators and barriers to implementing telemedicine for hypertension management [13], only focused on telemedicine, whereas our study examined all DHTs from the perspective of both patients and providers. Consistent with Mileski et al, we found that improved outcomes, increased patient knowledge and self-management, and cost savings were important facilitators of DHT adoption. Another systematic literature review by Gagnon et al [15] evaluated the factors influencing adoption of DHTs by health care professionals and some barriers reported in this review, such as the lack of organizational support and lack of reimbursement for providers, these were consistent with our study findings. Furthermore, most of the studies included in the review by Gagnon et al were conducted in large hospitals. In contrast, most studies in our review, 86% (31 out of 36 studies), were conducted in primary care settings. Additionally, Gagnon et al [15] examined DHTs across multiple diseases, whereas our review focused specifically on DHTs for hypertension management.

Multiple conceptual models exist to describe acceptance and usage of technology, such as Rogers diffusion of innovations theory [54], the technology acceptance model [55], and the unified theory of acceptance and use of technology (UTAUT) [56]. These models have been applied to describe the adoption of electronic health records and other forms of DHTs [57]. As a thematic analysis approach was used to identify new or emergent themes, we neither tied our analysis to a preexisting conceptual model nor sought to validate a preexisting conceptual model. However, it is worth noting that the themes that emerged from our analysis align with several of the constructs described in UTAUT. For example, the themes of clinical workflow integration and technology usability relate to the UTAUT construct of effort expectancy. Similarly, the theme of improvement in patient outcomes relates to the UTAUT construct of performance expectancy.

Future Implications

Lack of usability or ease of use was found to be a major barrier for both patients and providers in our review. Furthermore, lack of integration with clinical workflow was an important barrier for physicians. In the light of these findings, it is important that developers of DHTs should aim to improve the experience of both patients and providers through human-centered design thinking principles [58]. Such a process considers the needs and perspectives of all stakeholders during the product development cycle and implementation in a health care setting. With the right design, providers can interact with DHTs more easily to gain valuable insights on their patients' health, without compromising their existing workflow. In addition, successful implementation of DHTs in the clinical setting demands time and resources; new programs deploying DHTs should assess all the additional resources required for managing and coordinating care of patients to reduce the burden on providers.

Furthermore, providers often require hospital leadership to be supportive of a culture of innovation within their organization while weighing risks and benefits to patients and providers [38]. Therefore, organizational commitment to engaging providers at an early stage of DHT implementation by evaluating provider needs, identifying provider champions for implementing DHTs, and providing adequate training in the hospitals are critical to foster adoption.

Although not a prominent theme in this review, some studies show that the current health care policy and regulatory landscape are increasing pressure on health care organizations to provide lower-cost and higher-quality health care [59,60]. With tightening health care budgets, identifying long-term return on investment (ROI) on DHTs and establishing financial incentives through a clear reimbursement policy for providers are vital factors in increasing provider adoption. Therefore, future studies should incorporate discussions of implementation costs and ROI, in addition to examining clinical outcomes seen as a result of DHTs.

Limitations

First, as technology and policy are evolving at a rapid pace, certain barriers and facilitators that were identified in older articles may be less relevant today. Nevertheless, some facilitators and barriers are likely to remain constant over time, such as the critical importance of integration of DHTs into clinical workflow and technology usability. Second, reporting barriers and facilitators was not the primary aim of some of the studies included in this review. Thus, a portion of the data was collected from impressions reported in discussion sections of the published studies, which includes interpretations and speculations made by the researchers involved in the studies. Finally, some of the studies included in this review provided

little context on barriers and facilitators reported. In such instances, reviewers used their best judgement to determine whether the barriers or facilitators reported were best categorized as provider- or patient-related barriers or facilitators. Regardless of limitations, the themes in this review provide comprehensive evidence that could better inform and strengthen DHT development and implementation.

Conclusions

Our findings suggest that DHT adoption for hypertension is influenced by several important factors such as integration into the clinical workflow, usability, improvements in patient outcomes, and positive impact on the patient-provider relationship. Real-world testing and incorporating feedback from both patients and providers in designing technologies will improve their overall usability. Finally, to fully realize the potential of digitally enabled hypertension management, there is a greater need to validate these technologies to provide patients and providers with reliable and accurate information on both clinical outcomes and cost effectiveness.

Authors' Contributions

RP, NF, and AC were the primary reviewers and they conducted the thematic analysis in addition to writing the paper. RP, SA, KK, and KJ developed the review plan, process, and methodology. SA, KK, JF, CK, SL, and KJ contributed to the interpretation of results, manuscript review and editing, and provided advice and guidance throughout the review process and manuscript preparation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PubMed Strategy.
[PDF File (Adobe PDF File), 33KB - cardio_v3i1e11951_app1.pdf]

Multimedia Appendix 2
Summary of characteristics and authors' analysis of the 36 studies included in the review.
[PDF File (Adobe PDF File), 23KB - cardio_v3i1e11951_app2.pdf]

References

1. Morton K, Dennison L, May C, Murray E, Little P, McManus RJ, et al. Using digital interventions for self-management of chronic physical health conditions: A meta-ethnography review of published studies. Patient Educ Couns 2017 Dec;100(4):616-635 [FREE Full text] [doi: 10.1016/j.pec.2016.10.019] [Medline: 28029572]
2. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
3. American Medical Association. 2016. Digital Health Study: Physicians' motivations and requirements for adopting digital clinical tools URL: http://tinyurl.com/y67l9zyj [accessed 2019-02-19] [WebCite Cache ID 76J6gTFIV]
4. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. Circulation 2018 Mar 20;137(12):e67-e492. [doi: 10.1161/CIR.0000000000000558] [Medline: 29386200]
5. Merai R, Siegel C, Rakotz M, Basch P, Wright J, Wong B, DHSc, et al. CDC grand rounds: a public health approach to detect and control hypertension. MMWR Morb Mortal Wkly Rep 2016 Nov 18;65(45):1261-1264 [FREE Full text] [doi: 10.15585/mmwr.mm6545a3] [Medline: 27855138]
6. Serumaga B, Ross-Degnan D, Avery AJ, Elliott RA, Majumdar SR, Zhang F, et al. Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. Br Med J 2011 Jan 25;342:d108 [FREE Full text] [doi: 10.1136/bmj.d108] [Medline: 21266440]
7. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 04;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]

8. Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM. Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. Hypertension 2006 Mar;47(3):345-351. [doi: 10.1161/01.HYP.0000200702.76436.4b] [Medline: 16432045]

9. McKinstry B, Hanley J, Wild S, Pagliari C, Paterson M, Lewis S, et al. Telemonitoring based service redesign for the management of uncontrolled hypertension: multicentre randomised controlled trial. Br Med J 2013 May 24;346:f3030 [FREE Full text] [doi: 10.1136/bmj.f3030] [Medline: 23709583]

10. Pickering TG, Miller NH, Ogedegbe G, Krakoff LR, Artinian NT, Goff D, American Heart Association, American Society of Hypertension, Preventive Cardiovascular Nurses Association. Call to action on use and reimbursement for home blood pressure monitoring: executive summary: a joint scientific statement from the American Heart Association, American Society Of Hypertension, and Preventive Cardiovascular Nurses Association. Hypertension 2008 Jul;52(1):1-9. [doi: 10.1161/HYPERTENSIONAHA.107.189011] [Medline: 18497371]

11. Ward AM, Takahashi O, Stevens R, Heneghan C. Home measurement of blood pressure and cardiovascular disease: systematic review and meta-analysis of prospective studies. J Hypertens 2012 Mar;30(3):449-456. [doi: 10.1097/HJH.0b013e32834e4aed] [Medline: 22241136]

12. Minetaki K, Akematsu Y, Tsuji M. Effect of e-health on medical expenditures of outpatients with lifestyle-related diseases. Telemed J E Health 2011 Oct;17(8):591-595. [doi: 10.1089/tmj.2011.0019] [Medline: 21939380]

13. Mileski M, Kruse CS, Catalani J, Haderer T. Adopting telemedicine for the self-management of hypertension: systematic review. JMIR Med Inform 2017 Oct 24;5(4):e41 [FREE Full text] [doi: 10.2196/medinform.6603] [Medline: 29066424]

14. Logan AG, Dunai A, McIsaac WJ, Irvine MJ, Tisler A. Attitudes of primary care physicians and their patients about home blood pressure monitoring in Ontario. J Hypertens 2008 Mar;26(3):446-452. [doi: 10.1097/HJH.0b013e3282f2fdd4] [Medline: 18300854]

15. Gagnon M, Desmartis M, Labrecque M, Car J, Pagliari C, Pluye P, et al. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. J Med Syst 2012 Feb;36(1):241-277 [FREE Full text] [doi: 10.1007/s10916-010-9473-4] [Medline: 20703721]

16. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]

17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]

18. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

19. Grant RW, Pandiscio JC, Pajolek H, Woulfe A, Pelletier A, Kvedar J, et al. Implementation of a web-based tool for patient medication self-management: the Medication Self-titration Evaluation Programme (Med-STEP) for blood pressure control. Inform Prim Care 2012;20(1):57-67 [FREE Full text] [Medline: 23336836]

20. Eklind H. A hub clinician's perspective on supported self-care through technology. Prim Health Care 2017 Jan 27;27(1):20-25. [doi: 10.7748/phc.2017.e1147]

21. North F, Elrashidi MY, Ward WJ, Takahashi PY, Ebbert JO, Ytterberg KL, et al. Telemonitoring blood pressure by secure message on a patient portal: use, content, and outcomes. Telemed J E Health 2015 Aug;21(8):630-636. [doi: 10.1089/tmj.2014.0179] [Medline: 25885765]

22. Ideal Life. White paper: Hypertension URL: http://www.mtelehealth.com/pdf/studies/IdealLifeWhitePaperonHypertension.pdf [accessed 2019-02-19] [WebCite Cache ID 76J2MFGjg]

23. Neumann CL, Menne J, Rieken EM, Fischer N, Weber MH, Haller H, et al. Blood pressure telemonitoring is useful to achieve blood pressure control in inadequately treated patients with arterial hypertension. J Hum Hypertens 2011 Dec;25(12):732-738. [doi: 10.1038/jhh.2010.119] [Medline: 21228822]

24. Hanley J, Ure J, Pagliari C, Sheikh A, McKinstry B. Experiences of patients and professionals participating in the HITS home blood pressure telemonitoring trial: a qualitative study. BMJ Open 2013 May 28;3(5) [FREE Full text] [doi: 10.1136/bmjopen-2013-002671] [Medline: 23793649]

25. Mann M, Qiu M. Fraser Street Medical. 2015. Can we improvide myocardial protection during ischemic injury? Addressing hypertension care gaps by implementing an evidence based electronic medical record (EMR) hypertension dashboard URL: http://docplayer.net/48682040-Can-we-improve-myocardial-protection-during-ischemic-injury.html [WebCite Cache ID 76XQw6B3I]

26. Crowley MJ, Smith VA, Olsen MK, Danus S, Oddone EZ, Bosworth HB, et al. Treatment intensification in a hypertension telemanagement trial: clinical inertia or good clinical judgment? Hypertension 2011 Oct;58(4):552-558 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.174367] [Medline: 21844490]

27. Wang V, Smith VA, Bosworth HB, Oddone EZ, Olsen MK, McCant F, et al. Economic evaluation of telephone self-management interventions for blood pressure control. Am Heart J 2012 Jun;163(6):980-986. [doi: 10.1016/j.ahj.2012.03.016] [Medline: 22709750]

28. Wald JS, Businger A, Gandhi TK, Grant RW, Poon EG, Schnipper JL, et al. Implementing practice-linked pre-visit electronic journals in primary care: patient and physician use and satisfaction. J Am Med Inform Assoc 2010;17(5):502-506 [FREE Full text] [doi: 10.1136/jamia.2009.001362] [Medline: 20819852]

29. Shelley D, Tseng TY, Matthews AG, Wu D, Ferrari P, Cohen A, et al. Technology-driven intervention to improve hypertension outcomes in community health centers. Am J Manag Care 2011 Dec;17(12 Spec No):SP103-SP110 [FREE Full text] [Medline: 22216768]

30. Varis J, Karjalainen S, Korhonen K, Viigimaa M, Port K, Kantola I. Experiences of telemedicine-aided hypertension control in the follow-up of Finnish hypertensive patients. Telemed J E Health 2009 Oct;15(8):764-769. [doi: 10.1089/tmj.2009.0029] [Medline: 19780688]

31. Kim YN, Shin DG, Park S, Lee CH. Randomized clinical trial to assess the effectiveness of remote patient monitoring and physician care in reducing office blood pressure. Hypertens Res 2015 Jul;38(7):491-497. [doi: 10.1038/hr.2015.32] [Medline: 25787041]

32. Sublet ML, Courand PY, Bally S, Krummel T, Dimitrov Y, Brucker M, et al. 3C.07 Are the physicians reluctant to practice telemedicine in hypertension? J Hypertens 2015 Jun;33:e38-e39. [doi: 10.1097/01.hjh.0000467451.68858.df]

33. Frias J, Virdi N, Raja P, Kim Y, Savage G, Unger J, et al. Evaluation of a digital health offering to optimize blood pressure and lipid control in patients with uncontrolled hypertension and type 2 diabetes. Circulation 2016 Nov;134(Suppl_1):E [FREE Full text]

34. Kressin N, Long J, Glickman M, Bokhour B, Orner M, Clark C, et al. A brief, multifaceted, generic intervention to improve blood pressure control and reduce disparities had little effect. Ethn Dis 2016 Dec 21;26(1):27-36 [FREE Full text] [doi: 10.18865/ed.26.1.27] [Medline: 26843793]

35. Nilsson M, Rasmark U, Nordgren H, Hallberg P, Skönevik J, Westman G, et al. The physician at a distance: the use of videoconferencing in the treatment of patients with hypertension. J Telemed Telecare 2009;15(8):397-403. [doi: 10.1258/jtt.2009.090509] [Medline: 19948706]

36. Bernocchi P, Scalvini S, Bertacchini F, Rivadossi F, Muiesan M. Home based telemedicine intervention for patients with uncontrolled hypertension--a real life non-randomized study. BMC Med Inform Decis Mak 2014 Jun 12;14:52 [FREE Full text] [doi: 10.1186/1472-6947-14-52] [Medline: 24920046]

37. Wakefield BJ, Koopman RJ, Keplinger LE, Bomar M, Bernt B, Johanning JL, et al. Effect of home telemonitoring on glycemic and blood pressure control in primary care clinic patients with diabetes. Telemed J E Health 2014 Mar;20(3):199-205 [FREE Full text] [doi: 10.1089/tmj.2013.0151] [Medline: 24404819]

38. Millery M, Shelley D, Wu D, Ferrari P, Tseng T, Kopal H. Qualitative evaluation to explain success of multifaceted technology-driven hypertension intervention. Am J Manag Care 2011 Dec;17(12 Spec No):SP95-S102 [FREE Full text] [Medline: 22216774]

39. Bosworth HB, Olsen MK, Dudley T, Orr M, Goldstein MK, Datta SK, et al. Patient education and provider decision support to control blood pressure in primary care: a cluster randomized trial. Am Heart J 2009 Mar;157(3):450-456. [doi: 10.1016/j.ahj.2008.11.003] [Medline: 19249414]

40. Vedanthan R, Blank E, Tuikong N, Kamano J, Misoi L, Tulienge D, et al. Usability and feasibility of a tablet-based Decision-Support and Integrated Record-keeping (DESIRE) tool in the nurse management of hypertension in rural western Kenya. Int J Med Inform 2015 Mar;84(3):207-219 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.12.005] [Medline: 25612791]

41. Shaw RJ, Kaufman MA, Bosworth HB, Weiner BJ, Zullig LL, Lee SD, et al. Organizational factors associated with readiness to implement and translate a primary care based telemedicine behavioral program to improve blood pressure control: the HTN-IMPROVE study. Implement Sci 2013 Sep 08;8:106 [FREE Full text] [doi: 10.1186/1748-5908-8-106] [Medline: 24010683]

42. DeAlleaume L, Parnes B, Zittleman L, Sutter C, Chavez R, Bernstein J, et al. Success in the Achieving CARdiovascular Excellence in Colorado (A CARE) home blood pressure monitoring program: a report from the Shared Networks of Colorado Ambulatory Practices and Partners (SNOCAP). J Am Board Fam Med 2015;28(5):548-555 [FREE Full text] [doi: 10.3122/jabfm.2015.05.150024] [Medline: 26355126]

43. Kim H, Ahn ME, Choi YA, Choi EH, Kim DW, Shin SG, et al. Fifteen-year experience with telemedicine services in Gangwon province in Korea. Healthc Inform Res 2015 Oct;21(4):283-291 [FREE Full text] [doi: 10.4258/hir.2015.21.4.283] [Medline: 26618035]

44. O'Connor PJ, Magid D, Sperl-Hillen J, Price D, Asche S, Rush W, et al. Personalised physician learning intervention to improve hypertension and lipid control: randomised trial comparing two methods of physician profiling. BMJ Qual Saf 2014 Dec;23(12):1014-1022 [FREE Full text] [doi: 10.1136/bmjqs-2014-002807] [Medline: 25228778]

45. Holbrook A, Pullenayegum E, Thabane L, Troyan S, Foster G, Keshavjee K, et al. Shared electronic vascular risk decision support in primary care: computerization of Medical Practices for the Enhancement of Therapeutic Effectiveness (COMPETE III) randomized trial. Arch Intern Med 2011 Oct 24;171(19):1736-1744. [doi: 10.1001/archinternmed.2011.471] [Medline: 22025430]

46. Frias J, Virdi N, Raja P, Kim Y, Savage G, Osterberg L. Effectiveness of digital medicines to improve clinical outcomes in patients with uncontrolled hypertension and type 2 diabetes: prospective, open-label, cluster-randomized pilot clinical trial. J Med Internet Res 2017 Dec 11;19(7):e246 [FREE Full text] [doi: 10.2196/jmir.7833] [Medline: 28698169]

47. Abdullah A, Liew SM, Hanafi NS, Ng CJ, Lai PS, Chia YC, et al. What influences patients' acceptance of a blood pressure telemonitoring service in primary care? A qualitative study. Patient Prefer Adherence 2016;10:99-106 [FREE Full text] [doi: 10.2147/PPA.S94687] [Medline: 26869773]

48. Czaja SJ, Lee CC, Arana N, Nair SN, Sharit J. Use of a telehealth system by older adults with hypertension. J Telemed Telecare 2014 Jun;20(4):184-191 [FREE Full text] [doi: 10.1177/1357633X14533889] [Medline: 24803275]

49. Chen MJ, Chen KY, Chiang SJ, Lee JS, Ernest, Yu W, et al. E-010 the community-based telehealth care model for hypertension in Taipei. J Hypertens 2011;29:e12. [doi: 10.1097/01.hjh.0000408015.02905.f2]

50. Cottrell E, McMillan K, Chambers R. A cross-sectional survey and service evaluation of simple telehealth in primary care: what do patients think? BMJ Open 2012;2(6) [FREE Full text] [doi: 10.1136/bmjopen-2012-001392] [Medline: 23192242]

51. Bodenheimer T, Chen E, Bennett HD. Confronting the growing burden of chronic disease: can the US health care workforce do the job? Health Aff (Millwood) 2009;28(1):64-74. [doi: 10.1377/hlthaff.28.1.64] [Medline: 19124856]

52. Huff LS, Zittleman L, DeAlleaume L, Bernstein J, Chavez R, Sutte C, et al. What keeps patients from adhering to a home blood pressure program? J Am Board Fam Med 2011;24(4):370-379 [FREE Full text] [doi: 10.3122/jabfm.2011.04.100266] [Medline: 21737761]

53. Watson AJ, Singh K, Myint- U, Grant RW, Jethwani K, Murachver E, et al. Evaluating a web-based self-management program for employees with hypertension and prehypertension: a randomized clinical trial. Am Heart J 2012 Oct;164(4):625-631. [doi: 10.1016/j.ahj.2012.06.013] [Medline: 23067923]

54. Rogers EM. Diffusion of Innovations. New York: Free Press; 1995.

55. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Dec;27(3):425-427 [FREE Full text] [doi: 10.2307/30036540]

56. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 Dec;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]

57. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]

58. van Gemert-Pijnen L, Kelders SM, Kipp H, Sanderman R. EHealth Research, Theory and Development: A Multi-Disciplinary Approach. Abingdon, UK: Routledge; 2018.

59. Schur CL, Sutton JP. Physicians in medicare ACOs offer mixed views of model for health care cost and quality. Health Aff (Millwood) 2017 Dec 01;36(4):649-654. [doi: 10.1377/hlthaff.2016.1427] [Medline: 28373330]

60. Ryan AM, Krinsky S, Maurer KA, Dimick JB. Changes in hospital quality associated with hospital value-based purchasing. N Engl J Med 2017 Dec 15;376(24):2358-2366 [FREE Full text] [doi: 10.1056/NEJMsa1613412] [Medline: 28614675]

Abbreviations

AMA: American Medical Association
BP: blood pressure
CINAHL: Cumulative Index to Nursing and Allied Health Literature
DHT: digital health technology
EMBASE: Excerpta Medica database
MeSH: medical subject heading
RCT: randomized controlled trial
ROI: return on investment
UTAUT: unified theory of acceptance and use of technology

Original Paper

Achieving Rapid Blood Pressure Control With Digital Therapeutics: Retrospective Cohort and Machine Learning Study

Nicole L Guthrie[1*], MS; Mark A Berman[1*], MD; Katherine L Edwards[1*], MSN; Kevin J Appelbaum[1*], BSE; Sourav Dey[2*], PhD; Jason Carpenter[2*], MS; David M Eisenberg[3*], MD; David L Katz[1,4*], MD, MPH

[1]Better Therapeutics, San Francsico, CA, United States

[2]Manifold, Inc, Oakland, CA, United States

[3]Department of Nutrition, Harvard TH Chan School of Public Health, Harvard University, Boston, MA, United States

[4]Yale-Griffin Prevention Research Center, Griffin Hospital, Yale School of Public Health, Derby, CT, United States

[*]all authors contributed equally

Corresponding Author:
Mark A Berman, MD
Better Therapeutics
445 Bush Street
Suite 300
San Francsico, CA, 94108
United States
Phone: 1 617 877 0327
Email: mark@bettertherapeutics.io

Abstract

Background: Behavioral therapies, such as electronic counseling and self-monitoring dispensed through mobile apps, have been shown to improve blood pressure, but the results vary and long-term engagement is a challenge. Machine learning is a rapidly advancing discipline that can be used to generate predictive and responsive models for the management and treatment of chronic conditions and shows potential for meaningfully improving outcomes.

Objective: The objectives of this retrospective analysis were to examine the effect of a novel digital therapeutic on blood pressure in adults with hypertension and to explore the ability of machine learning to predict participant completion of the intervention.

Methods: Participants with hypertension, who engaged with the digital intervention for at least 2 weeks and had paired blood pressure values, were identified from the intervention database. Participants were required to be ≥18 years old, reside in the United States, and own a smartphone. The digital intervention offers personalized behavior therapy, including goal setting, skill building, and self-monitoring. Participants reported blood pressure values at will, and changes were calculated using averages of baseline and final values for each participant. Machine learning was used to generate a model of participants who would complete the intervention. Random forest models were trained at days 1, 3, and 7 of the intervention, and the generalizability of the models was assessed using leave-one-out cross-validation.

Results: The primary cohort comprised 172 participants with hypertension, having paired blood pressure values, who were engaged with the intervention. Of the total, 86.1% participants were women, the mean age was 55.0 years (95% CI 53.7-56.2), baseline systolic blood pressure was 138.9 mmHg (95% CI 136.6-141.3), and diastolic was 86.2 mmHg (95% CI 84.8-87.7). Mean change was –11.5 mmHg for systolic blood pressure and –5.9 mmHg for diastolic blood pressure over a mean of 62.6 days (*P*<.001). Among participants with stage 2 hypertension, mean change was –17.6 mmHg for systolic blood pressure and –8.8 mmHg for diastolic blood pressure. Changes in blood pressure remained significant in a mixed-effects model accounting for the baseline systolic blood pressure, age, gender, and body mass index (*P*<.001). A total of 43% of the participants tracking their blood pressure at 12 weeks achieved the 2017 American College of Cardiology/American Heart Association definition of blood pressure control. The 7-day predictive model for intervention completion was trained on 427 participants, and the area under the receiver operating characteristic curve was .78.

Conclusions: Reductions in blood pressure were observed in adults with hypertension who used the digital therapeutic. The degree of blood pressure reduction was clinically meaningful and achieved rapidly by a majority of the studied participants.

XSL•FO
RenderX

Greater improvement was observed in participants with more severe hypertension at baseline. A successful proof of concept for using machine learning to predict intervention completion was presented.

KEYWORDS

Introduction

High blood pressure (BP), or hypertension, is the leading contributor of preventable death worldwide and based on the 2017 American College of Cardiology (ACC)/American Heart Association (AHA) guideline, it is prevalent among 45.6% of US adults [1]. This extraordinary prevalence is attributed, in part, to the omnipresent detrimental diet and lifestyle behaviors associated with hypertension [2,3].

The consequences of high BP have been appreciated since the 1930s, and an array of effective antihypertensive medications have been available for decades [4]. However, only half of those with hypertension have optimally controlled BP, and 16% have poorly controlled hypertension despite taking three or more antihypertensive medications [5,6].

In addition to pharmacotherapy, clinical guidelines in the United States and worldwide call for the initiation of behavioral therapy focused on lifestyle for all patients with hypertension, because it is known that lifestyle changes can directly lower BP while simultaneously improving other cardiovascular risk factors without the side effects of pharmacotherapy [1,7]. However, there is also widespread appreciation that the current health care system is unable to deliver behavioral therapies that predictably lead to sustained lifestyle changes among the massive volume of patients who need it [8,9].

As a part of a global call to address a worsening pandemic, technology companies have been asked to contribute innovative solutions that enhance BP control and reduce the burden of care on primary care systems [10,11]. In particular, digital interventions designed to treat chronic diseases, known as digital therapeutics, can be paired with remote monitoring devices to create novel means of delivering effective and highly accessible care. These same interventions can simultaneously monitor outcomes, as recent evidence demonstrated the validity and utility of BP monitoring at home [12,13].

There are numerous commercially available apps designed to aid BP management, especially BP-tracking apps, but very few of them are multicomponent behavioral interventions designed to treat hypertension and have been clinically evaluated [14-17]. The widespread availability of mobile health apps, and the difficulty patients and clinicians have in distinguishing between them, warrants more rigorous study and vetting [18].

The use of machine learning, a branch of artificial intelligence that aims to make sense of patterns within large datasets, offers the potential to further increase the effectiveness of digital interventions. For example, it can be used to predict the likelihood of a specific clinical outcome based on an individual's unique pattern of use of the multiple components that make up a digital intervention during the course of treatment. This predictive ability holds great promise in developing interventions that are precisely targeted to the individual for optimal effectiveness. It has been argued that improving our ability to target treatment to individual patients begins by identifying and addressing unique subgroups through advanced analytic techniques like machine learning and may be the best path forward to enact precision medicine [19,20].

Mobile apps are well situated to use machine learning, because they are continuously collecting engagement and biometric data and can be programmed to change the delivery of treatment in response to outputs of machine learning algorithms. Although the application of machine learning to mobile apps holds great promise, it has not been widely applied to mobile apps targeting the root causes of hypertension.

The digital intervention assessed in this paper was developed by Better Therapeutics LLC (San Francisco, CA), a developer of prescription digital therapeutics for the treatment of cardiometabolic diseases. The goal of this article is to provide a retrospective analysis of the effectiveness of digital therapeutics in delivering behavioral therapy to patients with hypertension, resulting in a reduction of BP. In addition, a proof of concept for the use of machine learning to predict intervention completion in a manner that allows for personalized, real-time treatment plan adjustments is presented.

Methods

Digital Intervention

The digital intervention integrates a mobile medical app that delivers behavioral therapy with the support of a remote multidisciplinary care team. The mobile app delivers a personalized behavior change intervention including tools for goal setting, skill building, self-monitoring, biometric tracking, and behavioral feedback. The intervention is designed to support the participant's daily efforts to reduce BP and improve overall cardiometabolic function by facilitating behavioral changes, such as planning and self-monitoring, that increase physical activity and change dietary pattern to one that is predominately made up of whole grains, fruits, vegetables, beans, legumes, nuts, and seeds. These targeted changes are consistent with well-established clinical guidelines [1,7]. Further, the app uses artificial intelligence to provide feedback and support during the intervention to enhance adherence to behavioral therapy and increase participants' self-efficacy to make and sustain behavioral changes. Use of the app is coupled with scheduled person-to-person health coaching by phone over a 12-week

treatment period. Program completion was defined as ongoing use of the intervention in week 10 or later.

The app was designed for daily use, with a typical interaction beginning in a conversational interface that prompts a participant to report their progress toward individualized behavioral goals, such as the number of plant-based meals and minutes of physical activity completed that day as well as any biometrics, such as BP or weight, that were recorded. The participant receives feedback based on the data collected and is then prompted to engage in one or more behavioral exercises. For example, they may be prompted to respond to a question from their coach to self-reflect on opportunities and barriers to meeting their weekly goals or begin a skill-learning exercise that challenges the participant to try a new method for preparing vegetables or a different strategy for incorporating exercise into their day.

Intervention participants were recruited directly through Facebook and employer-sponsored advertisements. The intervention was advertised as a 12-week program for adults who wanted to improve hypertension, type 2 diabetes, or hyperlipidemia. All enrollees who self-identified as hypertensive were provided the option to receive a free Omron 7 Series Upper Arm Blood Pressure Monitor (Omron Healthcare, Inc, Kyoto, Japan) for use throughout the intervention and to keep after the study ended. The intervention was available to individuals at no cost.

Intervention Participants

The intervention database was searched to identify participants with a starting BP value in the hypertensive range ($\geq$130/80 mmHg), as defined by the 2017 ACC/AHA guideline [1], as well as participants with elevated BP who reported using antihypertensive medication. From this group of participants, analysis cohorts were identified based on engagement with the intervention (Figure 1). The intervention days were counted from day 0 (account created), with day 1 being the first full day of access to the digital intervention.

Figure 1. Participant flow chart. BP: blood pressure.
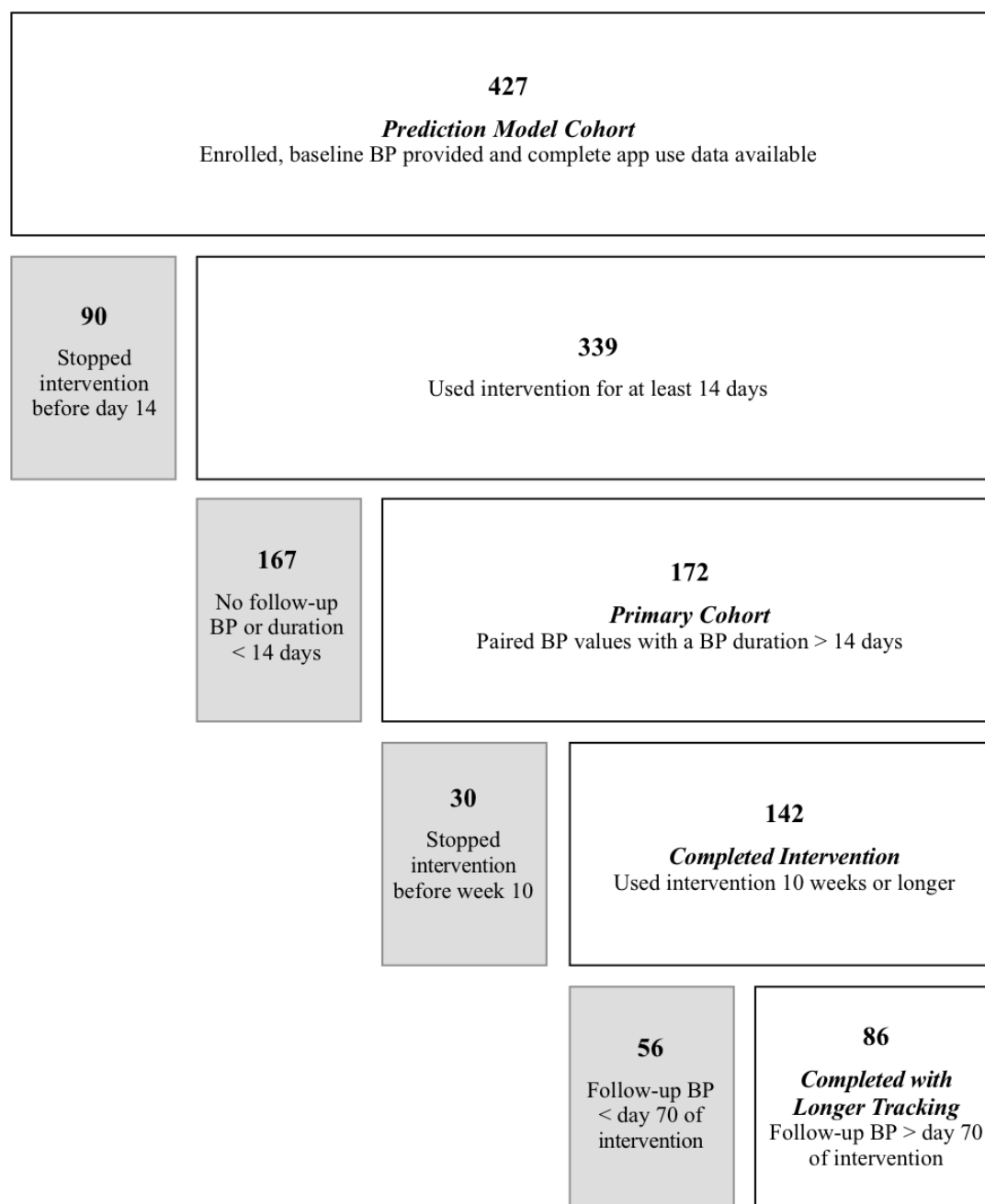
The *Primary* cohort included participants with hypertension, who provided follow-up BP values at least 14 days after the baseline. The *Completed Intervention* cohort included those who showed activity in the app in or after week 10 of the intervention. Participants were further categorized into the *Completed with Longer Tracking* cohort if, in addition to completing the intervention, their follow-up BP value was reported on or after day 70 of the intervention. The criteria for the analysis cohorts were defined prior to completing the analysis in order to explore the relationship between app engagement (*Completed Intervention*) and self-tracking (*Completed with Longer Tracking*) at the primary end point of changes in BP.

Data were identified by a unique numeric identification assigned by the system at registration; exported data included no personal identifiers. This retrospective analysis was approved and overseen by the Quorum Review Institutional Review Board [21], and a waiver of informed consent was granted for this retrospective analysis.

Measures

Demographics

Participants eligible for enrollment were adults, aged 18 years or older, were living in the United States, and had a smartphone with Android or Apple operating system to access the intervention app. Within the app, participants reported their age, gender, height, weight, medical history, state of residence, and current prescription medications.

Blood Pressure

Participants recorded BP readings in the app at will. Each reported measurement included a value for systolic BP (SBP), diastolic BP (DBP), and date and time of measurement. Baseline and follow-up values were calculated by taking an average of all available values in 7-day intervals. Day 1 of the intervention was set as the anchor day for the baseline interval, and all values reported within the following 6 days were included in the mean. The follow-up anchor date was set as the date of the last BP value reported, and all values reported in the 6 days prior to the anchor were included in the mean. The number of days between the baseline and follow-up BP was considered the duration of change. The mean change was calculated by subtracting the mean baseline value from the mean follow-up value. BP categorization was based on published guidelines (defining elevated as 120-129/<80 mmHg, stage I hypertension as 130-139/80-89 mmHg, and stage II hypertension as ≥140/≥90 mmHg) [1].

Weight

In addition to reporting weight at the time of enrollment, participants had the option to track their weight using their own home scale and record it in the app at will. Body mass index was calculated by dividing the weight in kilograms by the height in meters squared.

Predictive Modeling

We used machine learning to generate a model to predict whether someone would complete the intervention, applying the same criteria as defined above for the *Completed Intervention* cohort. A random forest model was trained on 427 participants with complete app use data available. The random forest model was selected because it reduces overfitting of a model by taking the average of many decision trees, which is important in small data sets [22]. A supervised classification algorithm was used, since the response variable *Completed Intervention*, is binary.

The random forest model was trained with 250 trees and a minimum of 3 samples per leaf node, as determined by hyperparameter optimization. The model included 19 features

that can be grouped as follows: (1) Engagement: These features were actions related to use of the intervention, such as the count of plant-based meals logged, skill-building modules completed, or health coaching calls completed. (2) Sociomarkers: These were indicators of social conditions that an individual is exposed to or surrounded by, which can be correlated with the presence or severity of a health state, such as zone improvement plan code or availability of health care [23]. Our model incorporated the novel sociomarkers' operating system (Android or Apple operating system) and email domain (Gmail, Yahoo, Hotmail, or Other), because we hypothesized that these sociomarkers may have predictive power. (3) Biometrics: These included the count of BP values reported, the baseline BP value, the count of weight values reported, and the percentage of weight loss. A list of all features included in the model is presented in Figure 2.

We trained random forest models on days 1, 3, and 7 of the intervention, with day 1 being the first full day of intervention engagement after the participant signs up. Development of training models at differing time points from the start day allowed us to explore the duration of engagement needed before predictive capacity emerged. For each model, only the data collected up to that day were used as features in the model. For example, in the day 3 model, we only used the engagement information collected in the first 3 days, and not beyond. For each model, the final response variable was the same—whether the patient completed the intervention. The training of the model includes a series of decision trees that evaluates data from the engagement features, sociomarkers, and biometrics, in relation to the response variable of interest—intervention completion.

We assessed generalization performance of the model by using leave-one-out cross-validation, which is a common technique for assessing model performance in samples of this size [24-26]. To this end, we trained the model on N–1 samples of the data and made a prediction on the one sample that was left out. This produces "out of sample" predictions for all N samples. These N predictions were pooled to compute various classification metrics, like the receiver operator characteristic (ROC), the area under the curve (AUC) of the ROC, and a confusion matrix of true versus predicted labels [27].

Figure 2. Shapley values illustrate which factors contribute most to an increased likelihood of completion (or noncompletion). Each dot represents the value of one individual participant for the feature component listed on the Y-axis. The value is represented by color (high vs low value) and by placement on the X-axis (amount of positive vs negative contribution to intervention completion). The feature list on the left is in order of contribution to the model (most to least). OS: operating system; SBP: systolic blood pressure; DBP: diastolic blood pressure; BP: blood pressure; BMI: body mass index; SHAP: Shapley Additive Explanation.
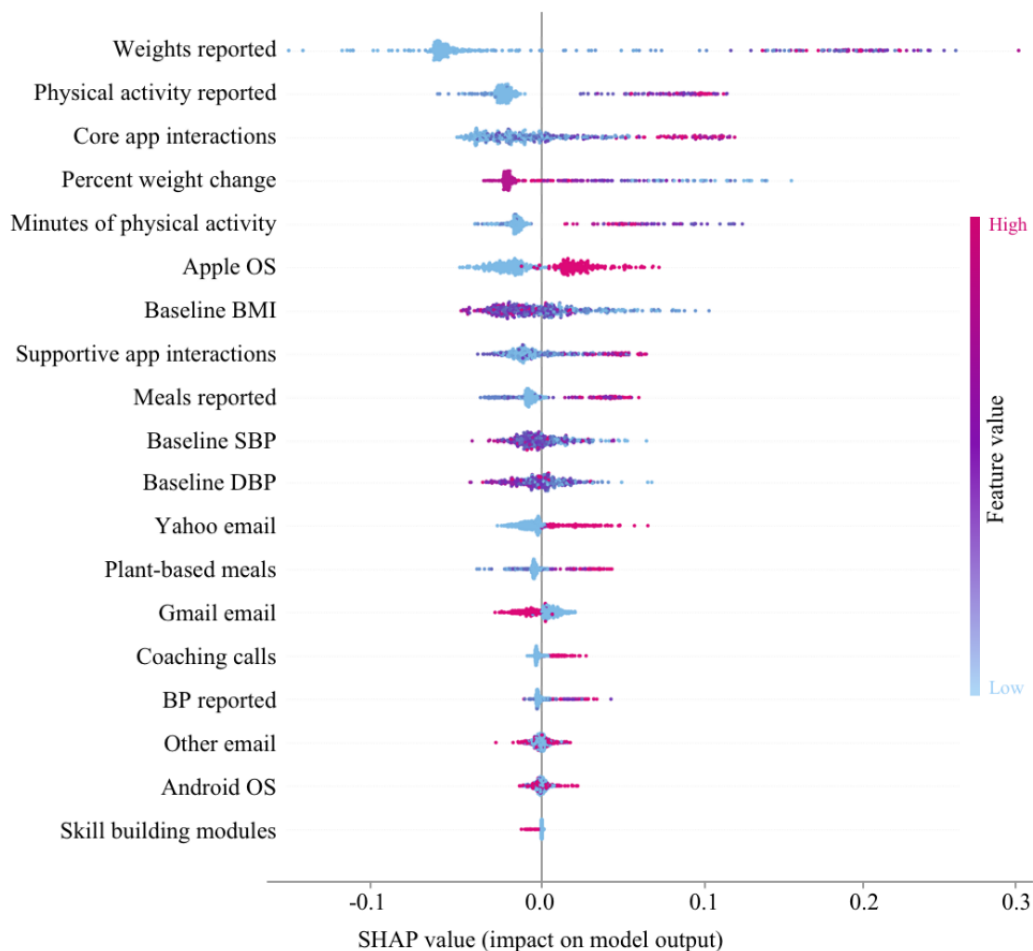
In addition, we used the Tree Shapley Additive Explanation (SHAP) algorithm [28], an explainable machine learning technique, on the random forest model to provide more interpretable predictions for each participant incorporated in the model. The Tree SHAP algorithm assigns each feature an importance value for every prediction. Each prediction begins at a base value, which is the expectation of the response variable (in our case, a probability between 0 and 1). Then, the SHAP values attribute to each feature the change in the expected model prediction when conditioning on that feature. The sum of the base value and all the additive feature attributions equal the final prediction probability.

All machine learning model development was performed using open-source packages in Python (Python Software Foundation, Wilmington, DE). The packages include but are not limited to Scikit-Learn, SHAP, Pandas, and Numpy.

Statistical Methods

Statistical analyses of changes in BP were performed using SAS software, version 9.4 (SAS Institute, Inc, Cary, NC). Change of continuous variables over time was analyzed using a two-tailed paired Student t test with alpha set at .05 and chi-square tests for differences in categorical variables. We used mixed-effects modeling to test the effects of baseline body mass index, baseline SBP, age, and gender on the mean change in BP.

Results

Intervention Participants

We identified 172 participants with hypertension (baseline BP≥130/80 mmHg or reported use of an antihypertensive medication) who engaged with the intervention for at least 2 weeks, reported a follow-up BP value, and were included in the *Primary* cohort. Demographics and baseline measurements for each cohort are presented in Table 1. There were no statistical differences in the baseline characteristics between those in the primary cohort and those in the two subgroups, as described above (ie, those who completed the intervention and those who completed and had a longer BP-tracking duration).

Blood Pressure

In the *Primary* cohort, 75.0% (129/172) of participants had a clinically meaningful improvement in BP (defined as a decrease of ≥5 mmHg in SBP or ≥2.5 mmHg in DBP). The mean change from baseline to last follow-up reported was –11.5 mmHg for SBP and –5.9 mmHg for DBP, with a mean duration between values of approximately 9 weeks (62.6 days). An improvement

of one BP category, as defined by ACC/AHA clinical practice guidelines [1], was seen in 51.7% (89/172) of the primary cohort. The changes between the end values of SBP and DBP were found to be significantly different from the corresponding baseline values ($P<.001$). The difference remained significant in the mixed-effects model accounting for the baseline SBP, age, gender, and body mass index ($P<.001$). The percent weight change was not found to correlate with changes in SBP ($P=.53$) or DBP ($P=.12$). Table 2 presents the changes in BP for the three analysis cohorts.

The mean duration between the baseline and final BP values for the *Completed Intervention* cohort was 10 weeks, with 74.7% (106/142) showing a meaningful improvement in BP and 22.5% (32/142) achieving a normal BP (BP<120/80 mmHg). The mean duration for the *Completed with Longer Tracking* cohort was 12.3 weeks, with 82.6% (71/86) of participants showing meaningful improvement and 26.7% (23/86) ending the

intervention with BP in the normal range. The percentage of participants with meaningful improvements in BP was higher in this cohort than the *Primary* cohort ($P=.02$).

Figure 3 contrasts the improvements seen in participants with stage I and stage II hypertension. Participants with stage I hypertension (n=76) saw a decrease of 5.4 mmHg (95% CI –7.4 to –3.3) in SBP and a decrease of 3.8 mmHg (95% CI –5.3 to –2.3) in DBP. Participants with stage II hypertension (n=84) observed a larger decrease in BP values, with SBP decreasing by 17.6 mmHg (95% CI –21.2 to –14.1) and DBP decreasing by 8.8 mmHg (95% CI –11.3 to –6.4).

Mean weekly SBP values from the *Completed with Longer Tracking* cohort were used to explore the rate of BP change (Figure 4). Although the mean BP continued to decline throughout the intervention period, the rate of decline was approximately 5 times greater in the first 6 weeks than in the following 6 weeks.

Table 1. Sample characteristics at baseline by intervention completion.

Participant characteristics	Primary cohort (N=172)	Completed intervention (N=142)	Completed with longer tracking (N=86)	P value[a]
Age[b] (years), mean (95% CI)	55.0 (53.7-56.2)	55.0 (53.7-56.4)	55.1 (53.2-56.9)	.87
Body mass index (kg/m^2), mean (95% CI)	35.3 (34.0-36.6)	34.9 (33.5-36.2)	34.3 (32.7-35.9)	.15
Female gender, n (%)	148 (86.1)	125 (88.0)	75 (87.2)	.66
Geographic distribution[c] (number of US states)	28	28	23	.77
Systolic BP[d] (mmHg), mean (95% CI)	138.9 (136.6-141.3)	138.6 (136.0-141.2)	138.1 (134.7-141.5)	.49
Diastolic BP (mmHg), mean (95% CI)	86.2 (84.8-87.7)	86.1 (84.5-87.7)	87.4 (85.3-89.4)	.12
Number of BP medications, mean (95% CI)	1.3 (1.2-1.5)	1.3 (1.1-1.5)	1.2 (0.96-1.5)	.12

[a]P value comparing the primary cohort to participants completing the intervention with longer tracking.

[b]Age was not available for 5 participants.

[c]US state data were not available for 50 participants.

[d]BP: blood pressure.

Table 2. Change in blood pressure across sample cohorts.

Measures	Primary cohort (N=172)	Completed intervention (N=142)	Completed and longer tracking (N=86)
Systolic BP[a] change (mmHg), mean (95% CI)	–11.5 (–13.7 to –9.3)	–11.2 (–13.6 to –8.8)	–12.7 (–16.0 to –9.5)
Diastolic BP change (mmHg), mean (95% CI)	–5.9 (–7.3 to –4.4)	–5.8 (–7.5 to –4.1)	–7.4 (–9.7 to –5.1)
BP duration (days), mean (95% CI)	62.6 (58.4 to 66.8)	68.5 (64.1 to 72.8)	86.5 (84.2 to 88.7)
Number of average weekly BP readings[b], mean (95% CI)	2.7 (2.4 to 3.1)	2.8 (2.4 to 3.2)	3.2 (2.6 to 3.7)
Meaningful changes in BP, n (%)	129 (75.0)	106 (74.7)	71 (82.6)
Follow-up BP average<140/90 mmHg, n (%)	132 (76.7)	108 (76.1)	69 (80.2)
Follow-up BP average<130/80 mmHg, n (%)	63 (36.6)	52 (36.6)	37 (43.0)
Follow-up BP average<120/80 mmHg, n (%)	39 (22.7)	32 (22.5)	23 (26.7)

[a]BP: blood pressure.

[b]Meaningful change is defined as a minimum decrease of 5 points in systolic blood pressure or 2.5 points in diastolic blood pressure.

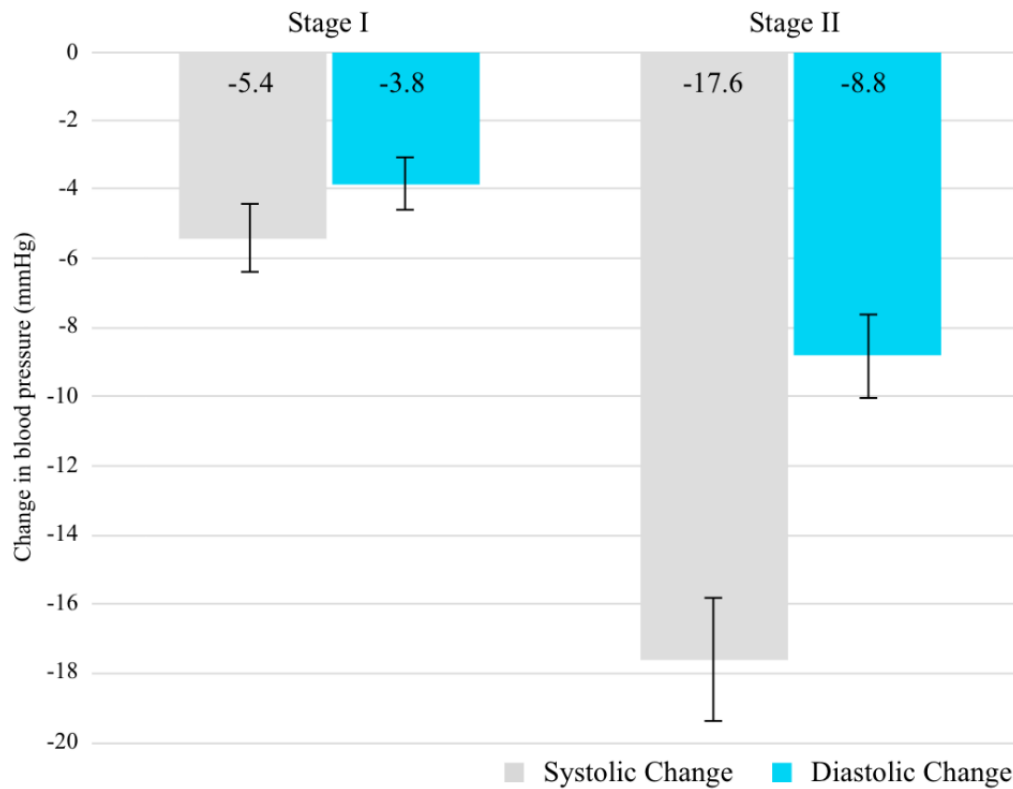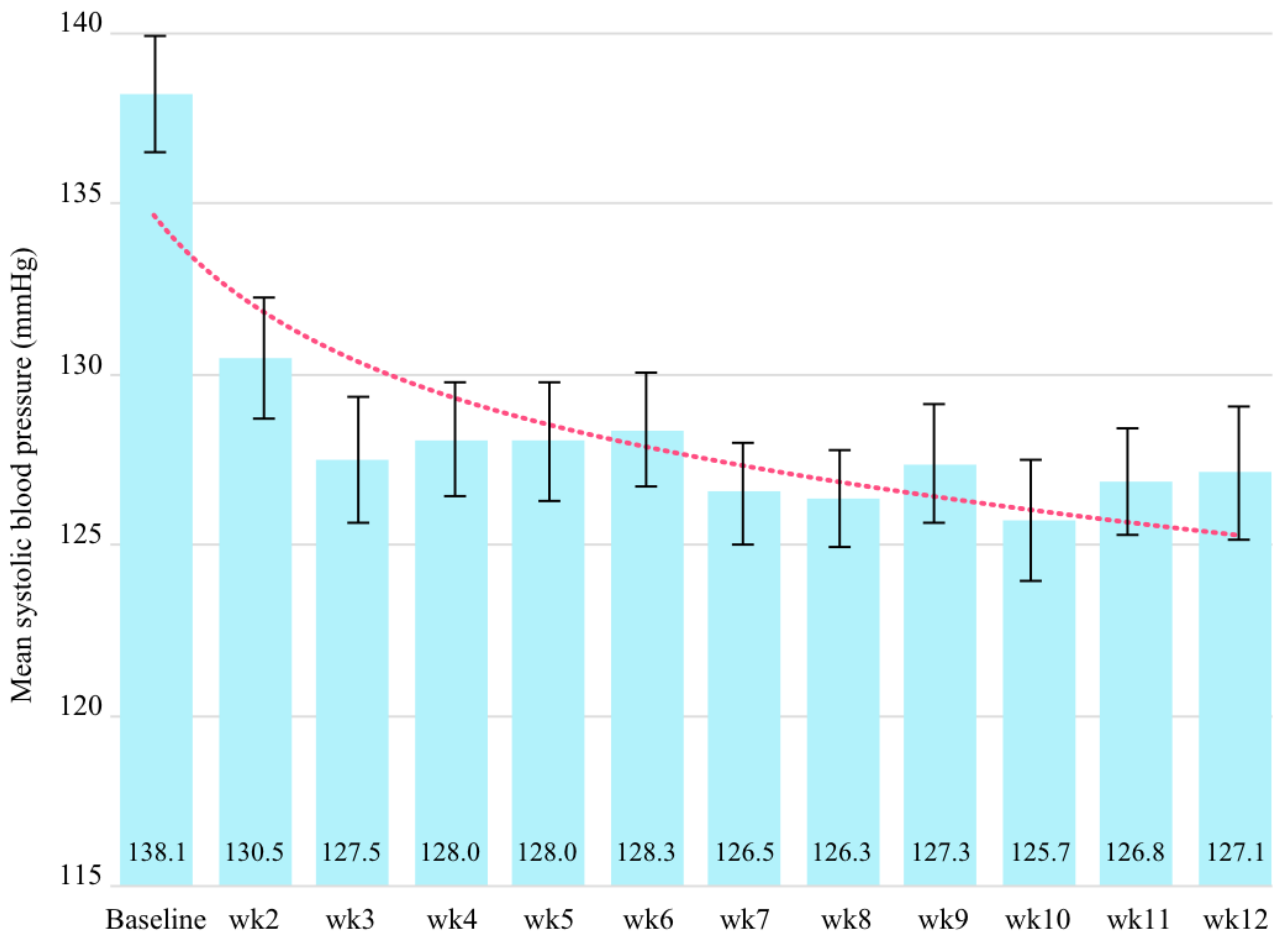**Figure 3.** Observed change in participant blood pressure by baseline category.

Figure 4. Mean systolic blood pressure over time in the Completed with Longer Tracking cohort. Plot of mean systolic blood pressure per intervention week, with SE bars. The sample size of weekly means varied from 42 to 86 participants.

Engagement

In the *Primary* cohort, 94.2% (162/172) of the participants were using the app after 6 weeks (mid-intervention) and 82.6% (142/172) of the participants completed the intervention. Full app use data were available for 94.8% (163/172) of the participants in the *Primary* cohort. Use of the app was defined as active use of any feature, excluding the act of a login or logout. Total distinct app engagements averaged 12.2 per day (95% CI 10.9-13.4), and the average number of calls completed with an intervention program health coach was 3.4 (95% CI 3.1-3.7).

Predictive Modeling

The random forest model was trained on 427 participants (*Prediction Model* cohort in Figure 1). The resultant ROC curve and AUC for days 1, 3, and 7 models showed that the model performs better as more days of data are used (Figure 5).

Performance of the day 7 model was examined with a target false positive rate of 25%. The nearest point on the ROC curve to the desired false positive rate was at 26%. At this point in the curve, we observed a sensitivity or true positive rate of 70%

and a specificity or true negative rate of 74%. The observed misclassification rate or error rate was 27%. The positive predictive value was 56%, and the negative predictive value was 84%.

The Tree SHAP algorithm was applied to ascertain which model features best predicted intervention completion in the entire analyzed population (Figure 2). The results indicate that, on an average, early engagements directly related to intended behavioral interactions and changes (eg, self-monitoring of biometrics, completing supportive and core app interactions, or reporting more exercise) are most predictive of intervention completion; sociomarkers (eg, Android vs iPhone use) are also predictive but to a lesser degree.

To illustrate how the machine learning model can convey both a completion probability and the contribution of each feature to that computed probability for a single participant, a plot of the SHAP values for a random participant at day 7 is shown in Figure 6. In this example, the participant's tracking of BP and baseline body mass index contributes to a higher probability of completion, but this is partially counteracted by the lack of reporting in exercise and relatively low tracking of weight.

Figure 5. Receiver operating characteristic curves for predictive models of days 1, 3, and 7. ROC: receiver operating characteristic; AUC: area under the curve.
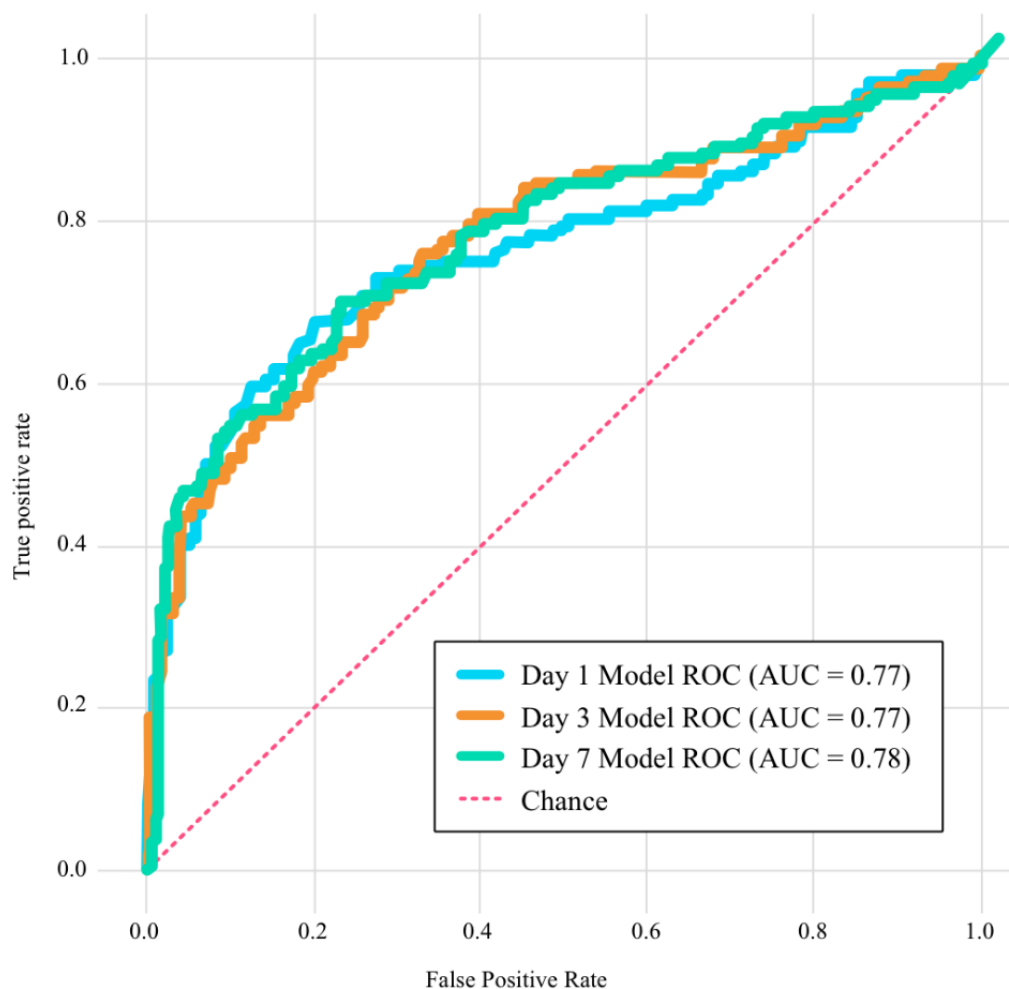
Figure 6. The day 7 probability of intervention completion for one participant, represented by the Shapley Additive Explanation Force Plot. Feature use contributing to a higher probability of intervention completion is shown in red, along with the size of the feature's contribution. Feature use contributing to a lower probability of intervention completion is shown in blue. DBP: diastolic blood pressure; BP: blood pressure; BMI: body mass index.
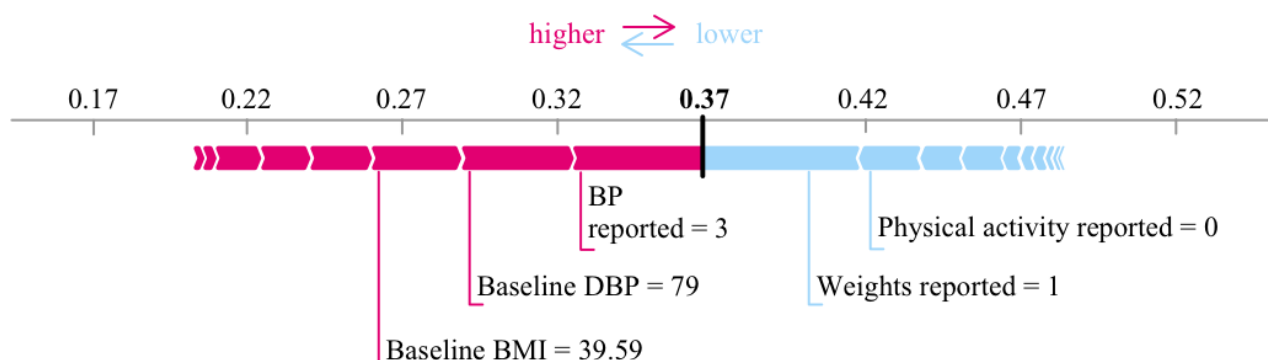
Discussion

Principal Findings

The digitally delivered intervention resulted in meaningful reductions in BP in adults. The majority of BP reduction was observed within the first 6 weeks of the intervention, indicating a rapid response to the digital intervention. By the end of the 12-week intervention, a high proportion of participants achieved BP control as per the ACC/AHA definition (43% in the *Completed with Longer Tracking* cohort) [1]. The greatest reductions were found in participants with stage 2 hypertension, with a mean SBP improvement of 17.6 mmHg. The BP-lowering effects observed were comparable or greater than those observed in other digitally delivered multicomponent interventions [14-16,24,29].

We did not find evidence that these improvements were the result of intensified medication therapy or that the BP reduction was due to weight change alone. This suggests that behavioral changes made during the intervention period account for much of the reduction in BP observed, and this observation is consistent with the effect sizes of other intensive behavioral or lifestyle interventions [29-32]. Importantly, the effects observed here are meaningfully greater than those observed by self-monitoring of BP alone, which suggests that multiple behaviors contribute to the effects [29,33]. For example, the 2017 meta-analysis conducted by Tucker et al showed that self-monitoring of BP was associated with changes of –3.2 mmHg in SBP and –1.6 mmHg in DBP between baseline and 12-month clinic measurements as compared to usual care [33]. In other analyses (data not shown), we did not find any correlation between the degree of self-tracking and BP change, nor did we find any difference between participants who were provided a home BP cuff and those who already had one.

A proof-of-concept analysis of a predictive model developed using machine learning demonstrated the ability to predict intervention completion after just one full day of engagement. The ability to predict intervention completion in a timely fashion is important for several reasons: (1) Given the typical patterns of apps use, there is likely a short time period during which an intervention adjustment can be made to increase completion rates. (2) Ongoing participation in the intervention is associated with a very high probability of achieving meaningful BP reductions. (3) Completion may be important for sustainment of behavioral changes and resulting outcomes beyond the intervention period.

This type of machine learning model can be implemented by choosing an operating point at which to make predictions. The operating point is chosen based on the balance of false versus true positives that it is expected to create. The prediction of likely to complete or not to complete the intervention can then be acted upon by leveraging SHAP values and creating an automated set of actions such as providing tailored feedback, reinforcement, warnings, and reprioritization of behavioral goals. The value of the model can then be studied in this context to see how it alters both completion rates and clinical outcomes.

This prediction methodology creates the opportunity for other exciting applications that may further improve the effectiveness of treatment. For example, the same methodology can be used to predict more direct measures of treatment success, such as a specific degree of BP improvement. Once a model that predicts clinical outcomes in the midst of treatment is validated, it can be used to alter the course of treatment with the intent of improving outcomes and patient experience.

Finally, machine learning allows us to explore discrete components of a digital intervention and the way they interact with participant characteristics. For example, in the current model, we found that the count of exercise sessions reported in the first week of the intervention was highly predictive of intervention completion. We also found that novel sociomarkers such as email domain or phone operating system had predictive capacity. For example, participants who used Yahoo email were more likely to complete the intervention than users of other email domains. It may be that Yahoo email use is a proxy for older age and other personality or socioeconomic features [34].

Limitations

A meaningful limitation of this retrospective analysis is the lack of a control group to evaluate the true effect size of this behavioral intervention. However, the effect size observed can be compared to similar study cohorts reported in the literature.

XSL•FO
RenderX

For example, in a recent 8-week study comparing metformin to placebo in nondiabetic adults with hypertension having similar demographic features and baseline BPs, the control group (n=49) had a mean improvement of 2.6 mmHg in SBP when measured in the clinic and a 0.7 mmHg mean increase when measured with 24-hour ambulatory BP monitoring [31]. Larger improvements of 6.0 mmHg in SBP were seen in the control group (n=131) in a 12-month study comparing the impact of electronic counseling on the standard of care for BP in adults with hypertension [29]. Participants of that study were recruited from the Heart and Stroke Canada website, and the authors hypothesized that this may have resulted in a study cohort of independently motivated participants, where participants assigned to the control group were more likely to take action with the standard of care recommendations. Improvements in that control group were clinically meaningful but are about half the size of those observed in our study cohort.

A limitation of our machine learning model is the size of the training dataset used, which typically correlates with the predictive strength of the model and limits the number of features that can be explored. However, it is encouraging to see that predictive power and feature importance can emerge from a relatively small dataset. This should encourage others to begin using machine learning models early, rather than waiting for massive datasets to accrue. The strength of the any machine learning model can be expected to improve over time as the training dataset grows.

Conclusions

Reductions in BP were observed among adults with hypertension who use the digital therapeutic studied here. The degree of BP reduction was clinically meaningful and achieved rapidly by a majority of participants studied. Greater improvement was observed in participants with more severe hypertension at baseline. A successful proof of concept for using machine learning to predict intervention completion after one day of app use was presented. Future research should examine the ability of treatment tailored in response to this model to further enhance outcomes. In addition, research is needed to assess the durability of outcomes following the intervention period, to identify subgroups and subgroup characteristics where the targeted intervention is most/least effective, and on the use of machine learning to predict clinical outcomes and modify treatment parameters during the course of treatment. The digital intervention should also be evaluated for its effectiveness in treating other chronic diseases that share the same root causes as hypertension.

Conflicts of Interest

NG, MB, KE, KA, and DK are employees and equity owners of Better Therapeutics, LLC (the study sponsor and intervention developer). DE, JC, and SD are independent paid scientific consultants of Better Therapeutics. JC was provided the raw data to perform all machine learning methods independently.

References

1. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, 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. J Am Coll Cardiol 2018 May 15;71(19):e127-e248. [doi: 10.1016/j.jacc.2017.11.006] [Medline: 29146535]

2. Fisher NDL, Curfman G. Hypertension—A Public Health Challenge of Global Proportions. JAMA 2018 Nov 06;320(17):1757. [doi: 10.1001/jama.2018.16760] [Medline: 30398584]

3. US Burden of Disease Collaborators, Mokdad AH, Ballestros K, Echko M, Glenn S, Olsen HE, et al. The State of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States. JAMA 2018 Dec 10;319(14):1444-1472. [doi: 10.1001/jama.2018.0158] [Medline: 29634829]

4. Gu Q, Dillon C, Burt V. Prescription drug use continues to increase: U.S. prescription drug data for 2007-2008. NCHS Data Brief 2010 Sep(42):1-8 [FREE Full text] [Medline: 20854747]

5. Daugherty SL, Powers JD, Magid DJ, Tavel HM, Masoudi FA, Margolis KL, et al. Incidence and Prognosis of Resistant Hypertension in Hypertensive Patients. Circulation 2012 Apr 03;125(13):1635-1642. [doi: 10.1161/CIRCULATIONAHA.111.068064]

6. Yoon SSS, Carroll MD, Fryar CD. Hypertension Prevalence and Control Among Adults: United States, 2011-2014. NCHS Data Brief 2015 Nov(220):1-8 [FREE Full text] [Medline: 26633197]

7. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, ESC Scientific Document Group. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur Heart J 2018 Sep 01;39(33):3021-3104. [doi: 10.1093/eurheartj/ehy339] [Medline: 30165516]

8. Qamar A, Braunwald E. Treatment of Hypertension. JAMA 2018 Nov 06;320(17):1751. [doi: 10.1001/jama.2018.16579]

9. Schroeder SA. Shattuck Lecture. We can do better--improving the health of the American people. N Engl J Med 2007 Sep 20;357(12):1221-1228. [doi: 10.1056/NEJMsa073350] [Medline: 17881753]

10. Kvedar J, Fogel A, Elenko E, Zohar D. Digital medicine's march on chronic disease. Nature biotechnology 2016;34(3):246. [doi: 10.1038/nbt.3495]

11. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]

12. McManus RJ, Mant J, Franssen M, Nickless A, Schwartz C, Hodgkinson J, et al. Efficacy of self-monitored blood pressure, with or without telemonitoring, for titration of antihypertensive medication (TASMINH4): an unmasked randomised controlled trial. The Lancet 2018 Mar;391(10124):949-959. [doi: 10.1016/S0140-6736(18)30309-X]

13. Melville S, Byrd JB. Out-of-Office Blood Pressure Monitoring in 2018. JAMA 2018 Nov 06;320(17):1805. [doi: 10.1001/jama.2018.14865]

14. 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]

15. 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]

16. Mao AY, Chen C, Magana C, Caballero BK, 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]

17. Liu S, Brooks D, Thomas SG, Eysenbach G, Nolan RP. Effectiveness of User- and Expert-Driven Web-based Hypertension Programs: an RCT. Am J Prev Med 2018 Dec;54(4):576-583. [doi: 10.1016/j.amepre.2018.01.009] [Medline: 29456025]

18. Bates DW, Landman A, Levine DM. Health Apps and Health Policy: What Is Needed? JAMA 2018 Dec 20;320(19):1975-1976. [doi: 10.1001/jama.2018.14378] [Medline: 30326025]

19. John HM. Managed Care Partners Internet. 2017 Jan 23. John Hopkins Medicine URL: https://www.hopkinsmedicine.org/news/publications/_docs/MCP1612009_JK_Managed%20Care%20Partners_Winter%202017-5.pdf [accessed 2018-11-28] [WebCite Cache ID 74Ry56KsW]

20. Shaban-Nejad A, Michalowski M, Buckeridge D. Health intelligence: how artificial intelligence transforms population and personalized health. npj Digital Med 2018 Oct 2;1(1). [doi: 10.1038/s41746-018-0058-9]

21. Quorum. URL: https://www.quorumreview.com/ [accessed 2017-12-07] [WebCite Cache ID 6vXFEq9l6]

22. Breiman L. Machine Learning. 2001. Random Forests URL: https://bit.csc.lsu.edu/~jianhua/random_forest.pdf [accessed 2019-01-17] [WebCite Cache ID 75V4CpQFE]

23. Shin EK, Mahajan R, Akbilgic O, Shaban-Nejad A. Sociomarkers and biomarkers: predictive modeling in identifying pediatric asthma patients at risk of hospital revisits. npj Digital Med 2018 Oct 2;1(1). [doi: 10.1038/s41746-018-0056-y]

24. Liu M, Chen X, Chen G, Cui Q, Yan G. A computational framework to infer human disease-associated long noncoding RNAs. PLoS One 2014 Jan;9(1):e84408 [FREE Full text] [doi: 10.1371/journal.pone.0084408] [Medline: 24392133]

25. Doytchinova I, Flower D. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007:7. [doi: 10.1186/1471-2105-8-4]

26. Moore RG, Brown AK, Miller MC, Skates S, Allard WJ, Verch T, et al. The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. Gynecol Oncol 2008 Feb;108(2):402-408. [doi: 10.1016/j.ygyno.2007.10.017] [Medline: 18061248]

27. Airola A, Pahikkala T, Waegeman W, De Baets B, Salakoski T. An experimental comparison of cross-validation techniques for estimating the area under the ROC curve. Computational Statistics & Data Analysis 2011 Apr;55(4):1828-1844. [doi: 10.1016/j.csda.2010.11.018]

28. Lundberg S, Erion GG, Lee SI. Cornell University. Consistent feature attribution for tree ensembles URL: https://arxiv.org/abs/1802.03888 [accessed 2019-01-17] [WebCite Cache ID 75V2n2HNS]

29. Nolan R, Feldman R, Dawes M, Kaczorowski J, Lynn H, Barr S, et al. Randomized Controlled Trial of E-Counseling for Hypertension: REACH. Circ Cardiovasc Qual Outcomes 2018 Jul;11(7):e004420. [doi: 10.1161/CIRCOUTCOMES.117.004420] [Medline: 30006474]

30. Ziv A, Vogel O, Keret D, Pintov S, Bodenstein E, Wolkomir K, et al. Comprehensive Approach to Lower Blood Pressure (CALM-BP): a randomized controlled trial of a multifactorial lifestyle intervention. J Hum Hypertens 2013 Oct;27(10):594-600 [FREE Full text] [doi: 10.1038/jhh.2013.29] [Medline: 23595161]

31. Júnior VC, Fuchs F, Schaan B, Moreira L, Fuchs S, Gus M. Effect of metformin on blood pressure in patients with hypertension: a randomized clinical trial. Endocrine 2019 Feb 21;63(2):252-258. [doi: 10.1007/s12020-018-1722-1] [Medline: 30132262]

32. 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 Apr 23;289(16). [doi: 10.1001/jama.289.16.2083]

33.   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 19;14(9):e1002389. [doi:
       10.1371/journal.pmed.1002389]
34.   Morning Consult: National Tracking Poll #170604. 2017 Jun. URL: https://morningconsult.com/wp-content/uploads/2017/
       06/170604_crosstabs_BRD_v3_TB.pdf [accessed 2018-12-05] [WebCite Cache ID 74QThFGwV]

Abbreviations

ACC: American College of Cardiology
AHA: American Heart Association
AUC: area under the curve
BP: blood pressure
DBP: diastolic blood pressure
ROC: receiver operator characteristic
SBP: systolic blood pressure
SHAP: Shapley Additive Explanation

Original Paper

Design of a Care Pathway for Preventive Blood Pressure Monitoring: Qualitative Study

Carlijn Geerse[1*], BSc; Cher van Slobbe[1*], BSc; Edda van Triet[1*], BSc; Lianne Simonse[1*], MSc, PhD

Product Innovation Management Department, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands
[*]all authors contributed equally

Corresponding Author:
Lianne Simonse, MSc, PhD
Product Innovation Management Department
Faculty of Industrial Design Engineering
Delft University of Technology
Landbergstraat 15
Delft, 2628 CE
Netherlands
Phone: 31 15 27 ext 89054
Email: L.W.L.Simonse@tudelft.nl

Abstract

Background: Electronic health (eHealth) services could provide a solution for monitoring the blood pressure of at-risk patients while also decreasing expensive doctor visits. However, a major barrier to their implementation is the lack of integration into organizations.

Objective: Our aim was to design a Care Pathway for monitoring the blood pressure of at-risk patients, in order to increase eHealth implementation in secondary preventive care.

Methods: A qualitative design study was used in this research. Data were collected by conducting visual mapping sessions including semistructured interviews with hypertension patients and doctors. The data were transcribed and coded and thereafter mapped into a Care Pathway.

Results: Four themes emerged from the results: (1) the current approach to blood pressure measuring has disadvantages, (2) risk and lifestyle factors of blood pressure measuring need to be considered, (3) there are certain influences of the at-home context on measuring blood pressure, and (4) new touchpoints between patients and health professionals need to be designed. These in-depth insights combined with the visualization of the current blood pressure process resulted in our Care Pathway design for monitoring the blood pressure of at-risk patients as secondary preventive care.

Conclusions: The Care Pathway guides the implementation of eHealth devices for blood pressure self-measurement. It showcases the pathway of at-risk patients and increases their involvement in managing their blood pressure. It serves as a basis for a new service using eHealth.

Introduction

Electronic Health and Relations Between Patient and Health Professional

The arrival of internet-enabled health (electronic health, eHealth), which ushered in large online health databases and platforms, caused a change in the attitude of health-conscious consumers and their relationship with health professionals. Health professionals are no longer the main source of health care information. Many consumers seek confirmation of their general practitioners' (GP) decisions on the Internet, enabling them to be more active in their health management [1-3]. For instance, in 2014, 73% of Dutch citizens obtained health-related information online [4]. The self-efficacy of patients has increased and the interaction between patient and GP has become more patient-centered [1]. The perception that patients should not be bothered with too much detailed information because they are unable to cope has become outdated [5]. According to the World Health Organization, patient

participation is the new paradigm in the global diffusion of eHealth—the use of information and communication technologies (ICT) for health. It enables increased self-managed care [6] and patient-centered care [7]. Self-management services can help patients understand their condition better and adjust their decision making accordingly [8]. They value data and information exchange more and want to be fully informed and continuously involved in the decision making in their treatment [3,9]. The overall challenge is how to embed these changes in new models of care—models that enable participation in the decision making enabled by eHealth and that share health care data, information, and insights into the relationship between patients and health professionals.

Integration Barriers for eHealth

To deal with higher demand for health care, social care, and social pensions from an ageing population and in the face of a shrinking workforce, the care models must be redesigned. Costs must be reduced without compromising the quality and accessibility of health care [10]. To achieve value-based health coverage, preventive eHealth services such as tracking diseases and monitoring public health are a necessity [11]. eHealth enabling early detection of diseases has a high positive impact on quality and costs. When symptoms are detected in an early stage, the likelihood of chronic diseases is reduced and high costs can be avoided [12]. Moreover, health monitoring as part of eHealth services could make it possible to avoid expensive hospital visits, and certain treatments of patients at risk can be done at home [13]. To date, however, there have been few successful implementations of eHealth services [14,15]. The slow adoption of services by users [16], the lack of collaboration between health care organizations and their systems [17], a shortage of funding [13], and the lack of care models to guide implementation are factors that obstruct the development and implementation of sustainable eHealth services [18]. This research therefore focuses on the design of a care model for eHealth services. We chose to study the monitoring of blood pressure, as it is an indicator for multiple high-cost diseases [19].

Blood Pressure Monitoring With eHealth for At-Risk Patients

Self-measured blood pressure devices have been around for many years but have not been extensively implemented in health services [20]. Early on, health professionals considered the results of these devices unreliable and most patients found them difficult to use [3]. Over the years, these eHealth products have improved considerably and are now trustworthy. A self-measured blood pressure device consists of an armband with sensors that measure one's blood pressure and heart rate and a monitor that presents the results and automatically transfers the data to an online app. This Web service provides the patient with data history and tracking overviews. For people with a higher risk or first symptoms of health problems, the use of self-measurement services could help detect illness or complications at an earlier stage. We refer to this group as at-risk patients. In the current health care system, the patient's blood pressure is measured during a GP consultation or at home. This is a reactive care service with one moment of measurement that

does not provide constant tracking and feedback on one's blood pressure. Researchers have found in favor of at-home measurements because it is difficult to take accurate blood pressure measurements in an unnatural environment [21]. The Dutch health care system organizes eHealth blood pressure monitoring with a preventive focus under a reimbursement-based financial model, in which disease complications and impairment are detected at an early stage, the so-called tertiary level of care provision [22]. The use of this model for eHealth services could shift towards secondary preventive care at the GP's office by helping detect complications in at-risk patients. eHealth can have major cost effects in avoiding a shift to primary care by specialists at the hospital. Regular blood pressure measuring, and thus earlier detection of hypertension, could reduce the likelihood of (exacerbation of) chronic diseases such as kidney failure [23]. A recent systematic review found that self-management of hypertension facilitated an increase in health, patient knowledge and involvement, greater cost-effectiveness, and a more accurate reading of results because the patient is in a natural environment [24]. The high blood pressure of this group of patients cannot be cured, but it can be managed. The estimated cost effects for the Netherlands are high, with 1.7 million people with chronic kidney disease, a number that is rising by 2000 per year [25]. Monitoring with eHealth could be a solution for monitoring patients' blood pressure while also decreasing expensive specialist visits.

Care Pathway Design to Overcome Barriers

The current approach to monitoring blood pressure of at-risk patients is used as tertiary preventive care, while eHealth services can make the shift to secondary preventive care. In order to make this shift, the current situation of monitoring blood pressure was researched. Communication and data streams between all the actors are important to design a Care Pathway for secondary preventive care. The major barrier that must be overcome in implementing preventive eHealth services is its lack of integration into organizations. The embedding of ICT requires care provider pathways that cross organizational boundaries. Care Pathways can be described as a concept for making patient-centered care operational [26] and gaining insight into how an organization can improve its services [27]. According to the European Pathway Association "the objective of a Care Pathway is to enhance the quality of care by improving patient outcomes, promoting patient safety, increasing patient satisfaction, and optimizing the use of resources" [28]. Studies have shown that Care Pathways improved communication between professionals [29] and clarified the division of roles and responsibilities [30]. With these qualities, a Care Pathway could help overcome some of the implementation barriers for preventive eHealth services. Therefore, in this paper, we concentrated on how to create a Care Pathway.

Research Question

The objective of our research is to design a Care Pathway for monitoring blood pressure of at-risk patients, focusing on the exchange of data and communication between the individuals involved. To do so, we will first define the current situation and use this to create a Care Pathway that focuses on clarifying how processes should be carried out and by whom in order to increase

the adoption of eHealth services [31]. The following research question was formulated: How can a Care Pathway to monitor blood pressure of "at-risk patients" be designed in order to increase the implementation of eHealth in secondary preventive care?

Methods

Study Design

For this research, a qualitative approach is used to explore this relatively new field. A qualitative research approach allows us to gain a better understanding of underlying opinions, reasons, and emotions [32]. A phenomenon study was chosen in order to unravel the participants' real experiences and understanding of the blood pressure monitoring service. According to Suter [33], the focus in a phenomenon study is on the essence of an experience, that is, on trying to understand the basic structure of that experience and interpreting the meaning it has for a person or group.

Data Collection

A visual mapping toolkit with an interview guide were used to organize generative sessions with the participants. As a result, a Care Pathway was co-created [34]. This was done to gain a better understanding of the data streams between the different individuals and locations where blood pressure is monitored. Three types of qualitative research data were obtained: audio recordings of the interviews, documentation of institutes, and visuals created in the visual mapping sessions.

Sampling

Purposeful sampling was used to identify and select individuals who had experiences with the studied phenomenon of eHealth [35]. For the design of a Care Pathway, it is especially important to involve different stakeholders in the development phase of the pathway [31]. In this research, we selected people diagnosed with high blood pressure and researched their current health

management and their interaction with health professionals. Seven interviews were conducted in total: one specialist, three patients, and three GPs. An overview of the participants can be found in Table 1.

All interviewed patients had been diagnosed with high blood pressure after several measurements taken by their GP and 24-hour measurement at home. For one day they carried a blood pressure device around that measured their blood pressure every 30 minutes for 24 hours. All were prescribed medicines and had a GP consultation at least once a year. Patient 1 is part of an eHealth monitoring program and therefore regularly measures his blood pressure by himself. Patients 2 and 3 obtain information about their blood pressure only during their yearly consultation.

Ethics

The research was performed according to the principles of the Helsinki Declaration and Nuremberg and was checked and reviewed by the human research ethics committee (HREC) of the Delft University of technology [36]. We did not involve vulnerable groups of children or patients over age 65 in our sample design. The participation in our research was voluntary. The methods and data used were checked and approved by the HREC and the data steward of the Industrial Design Engineering faculty.

Visual Mapping Session

To generate insights into the interactions between the actors and the data streams, a visual mapping toolkit was designed, as introduced and outlined in a design research protocol by Meeuwen, Walt Meijer, and Simonse [18]. This toolkit consisted of two-dimensional representations of different people and products involved in the medical process, pencils, markers and blank paper. An overview of the design toolkit can be found in Figure 1. A pilot session with test participants was done to improve the toolkit before conducting the sessions with the participants in this sample.

Table 1. Sample of participants.

Participants	Gender	Situation	Health service experience	eHealth – blood pressure experience
Patients				
Patient 1	Male	Kidney disease	±20 years	Yes
Patient 2	Female	Pregnancy hypertension	±20 years	No
Patient 3	Male	Stress due to work	±2 years	No
General practitioners (GPs)				
GP 1	Male	Working as GP in own practice	±40 years	No
GP 2	Male	Working as medical advisor at a health insurance company	±20 years	Yes
GP 3	Male	Working as GP, community practice (formerly a trauma specialist)	±3 years	No
Specialist				
Specialist 1	Male	Specialist Internal medicine	±11 years	Yes

XSL•FO

RenderX

Figure 1. Visual mapping toolkit.
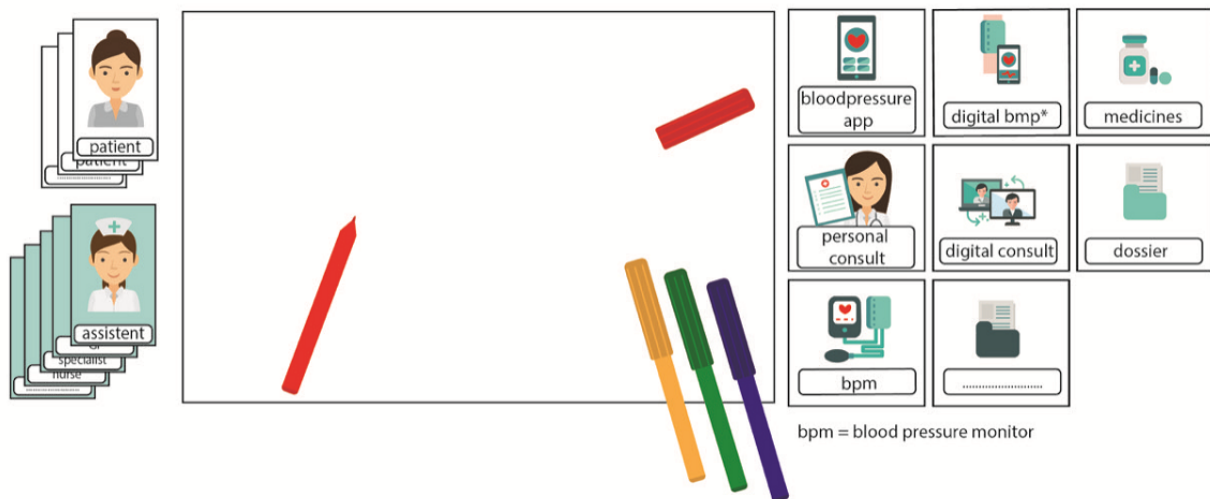
Six mapping sessions were performed with three patients and three GPs. The visual mapping session was built up in several levels since people's needs and values are often difficult to discuss [37]. The participants were first asked to visualize the current situation of blood pressure measuring and monitoring, and then to visualize a pathway for blood pressure monitoring at home. During the session, each participant was asked to reflect on their own needs and values [37], since they are the expert on their own experience area, including their interactions with the eHealth device and the ICT infrastructure [34]. The facilitator of the session played a guiding role during the visual mapping sessions and led the layering questioning. The visual mapping results were photographed at the end of each session in order to compare all the results.

Semistructured Interviews

The interview guide was constructed for the different actors, providing structure to the interviews while still maintaining the freedom to go into depth on certain topics [38]. The structure helped in analyzing the obtained data and finding patterns in it. All the interviews were audio recorded and transcribed for descriptive validity [39].

Data Analysis

Triangulation was used in the analysis by clustering the data from in-depth interviews, visual mapping sessions, and documentation [40]. First, initial coding identified important (groups of) words in the data, and these were labeled accordingly [41]. Second, with focused coding, codes were clustered and categorized to make a code book. Finally, multiple coding was used to avoid subjectivity in interpretation [42]. Subjectivity in interpretation was partly avoided by making use of investigator triangulation [40]. The outcomes of the mapping sessions were visually analyzed by comparing the different pathways of the participants. The similarities in the pathways have been considered and have led to the design of the current blood pressure monitoring situation and the new Care Pathway. The data analysis of the interviews served as a basis and check for the visual Care Pathway.

Results

Current Blood Pressure Monitoring Process

Before designing the Care Pathway, the current blood pressure process was visualized (Figure 2), drawn from the data obtained in the visual mapping sessions including the explanations in the interviews. In the current situation, the at-risk patient's blood pressure is monitored during a GP consultation as a physical check. In most cases, the GP assistant measures the patient's blood pressure and hands the results to the GP, who enters all the measured data manually into a health record system. If the patient's blood pressure is too high, a new appointment is made to repeat the measurement. If the blood pressure remains too high after several measurements have been taken within 3 months, the GP will prescribe medicines to lower and control the blood pressure. The pharmacist modifies the prescription if necessary. In exceptional cases, or when medicines do not lower the blood pressure, the patient will be referred to a specialist, who measures the blood pressure again.

On four themes that emerged from the data analysis, there is common agreement: (1) the current approach to blood pressure measuring has disadvantages, (2) risk and lifestyle factors of blood pressure measuring need to be taken into account, (3) there are certain influences of the at-home context on measuring blood pressure, and (4) new touchpoints between key actors need to be designed. Themes 1 and 2 relate to the current situation of blood pressure measuring and point out strengths and pitfalls. Theme 3 reveals important insights into self-measuring and monitoring blood pressure at home, and the perceived positive effects on patients' well-being, accuracy of measurement results, and health care costs and the negative effect of insecurity. Theme 4 extracts the new elements in the co-designed Care Pathway. The coding tree in Figure 3 shows the four themes with underlying categories and codes. All insights on these themes were considered in the design of the Care Pathway for blood pressure monitoring.

Figure 2. Current blood pressure process. GP: general practitioner.
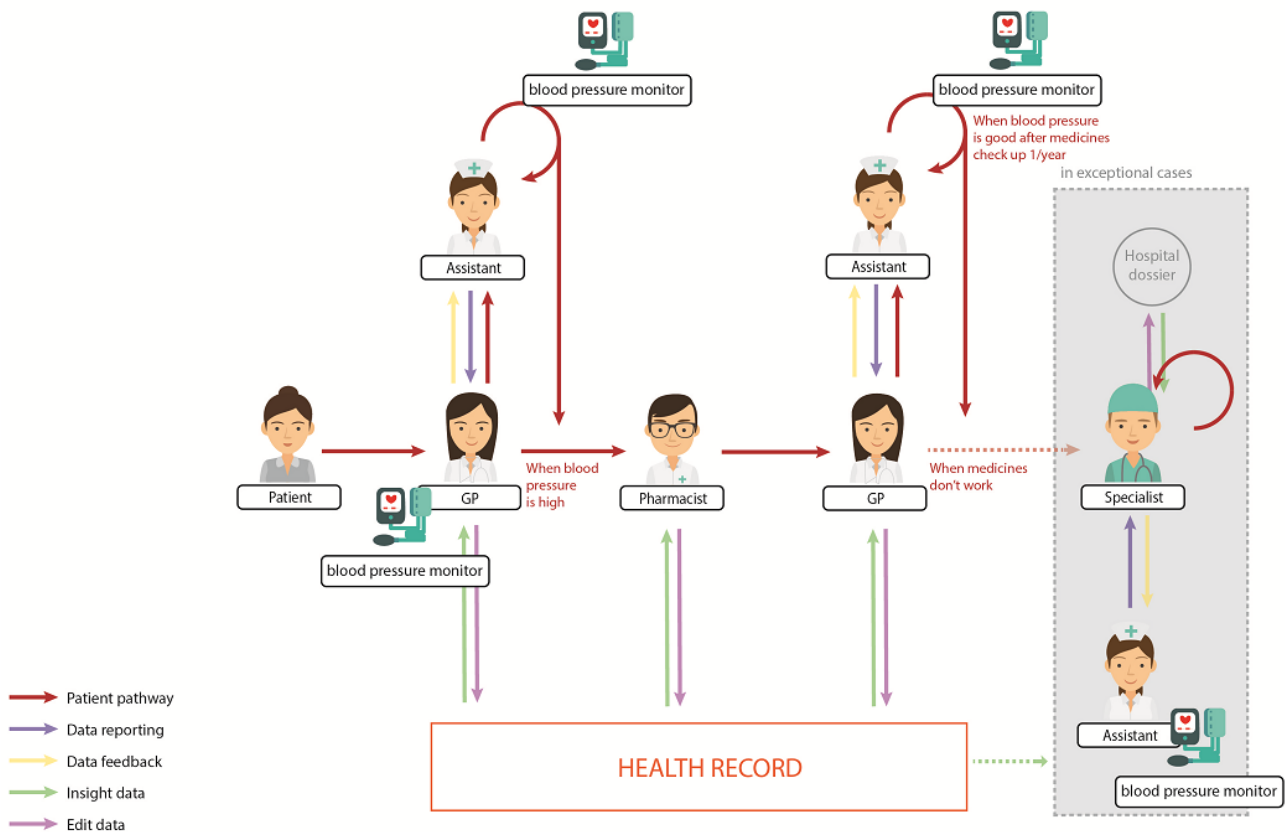
Theme 1: Disadvantages of Current Approach to Blood Pressure Measuring

One of the essential disadvantages of blood pressure measuring is that it is a snapshot of one measurement that varies over time. This drawback was quoted 27 times by participants (see the number 27 in brackets in Figure 2). Hereafter, we will refer in the text to the number of quotes as (# quotes). With the current approach to blood pressure measuring, there are little data about people's blood pressure, as it is measured only a few times a year during the GP consultation. Additionally, these measurements are not always accurate due to the "white coat effect" (7 quotes), meaning that some people become stressed when their blood pressure is being measured by a GP, resulting in an unrealistically high blood pressure. This effect can be reduced by taking multiple measurements over a certain period or by measuring in a different environment:

> *When I find a very high blood pressure, I just let the patients sit and relax and will do another measurement again. But you would prefer to have some more measurements over time.* [GP 2]

The GP has a prominent role in analyzing the results and deciding on the medication. In most cases the patient is not given insight into the blood pressure results. When the blood pressure of a patient is too high after several measurements, the GP often decides to switch to medication:

> *We usually treat at-risk patients with medicine. If one medicine does not work, you try a second one, then*

> *a third and maybe a fourth, all with an increasing dose.* [GP 1]

Furthermore, our participants validated that the current system is curative rather than preventive:

> *I do know that a lot of medication is prescribed, which may not be necessary. I am terrified how many medicines people swallow every day.* [Patient 2]

Participants suggested that this could be improved with the Care Pathway for blood pressure monitoring.

Theme 2: Risk and Lifestyle Factors on Blood Pressure Measuring

According to the GPs, blood pressure is considered a small piece of the complete health picture. Factors like smoking, diabetes, and high cholesterol in combination with blood pressure are important risk factors for heart and vascular diseases (13 quotes). Those risk factors are part of someone's lifestyle. The patients often stated that they have tried to adopt a healthier lifestyle after being diagnosed with high blood pressure (12 quotes).

> *The GP prescribed medicines but I didn't want to take them. So, I started doing sports andPatient 3losing weight, but getting stress under control was hard.* [Patient 3]

Both the specialist internal medicine and patients pointed out that lifestyle is hard to change, while lifestyle modification could also help lower blood pressure.

Figure 3. Coding tree with the analysis results of codes (number in brackets), categories and themes. GP: general practitioner.
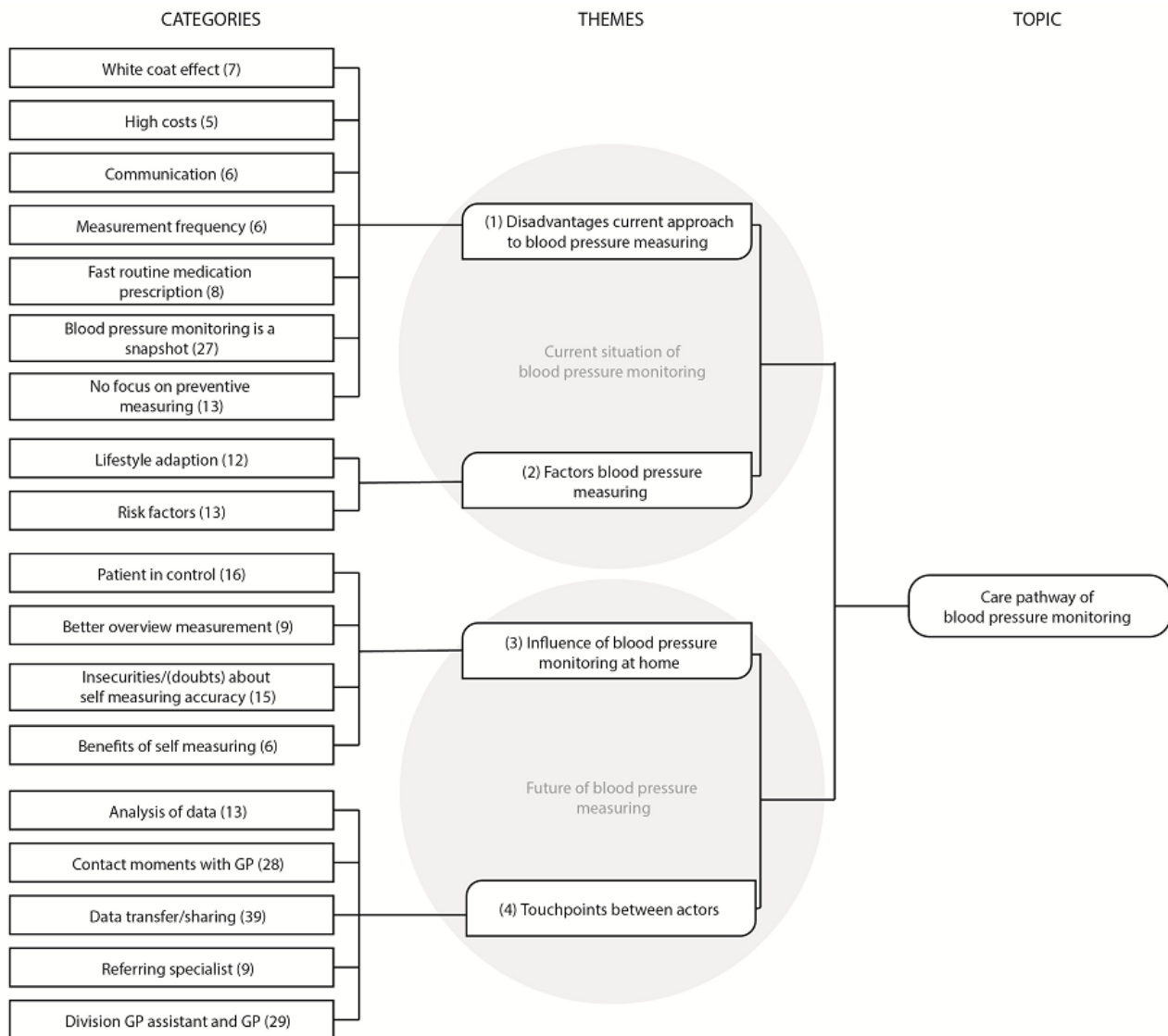
Theme 3: Influence of the At-Home Context on Blood Pressure Measuring

Measuring blood pressure at home enables patients to be in control of their own health care (16 quotes). Moreover, it gives a better overview and therefore more accurate measurements (9 quotes). Patients measure their blood pressure in a trusted environment at standard moments. This increases the reliability of the measurements. The benefit of self-measuring is that it reduces the influence of different context factors that can affect a person's blood pressure temporarily (6 quotes). Patients can decide when to measure and analyze their measurements themselves:

> *People enjoy measuring their blood pressure at home because it gives them power over their own health and they can try to influence it by looking at their lifestyle: if I eat less salt, does my blood pressure drop?* [GP 1]

We also observed a difference in attitude among the participants towards blood pressure measuring. GP 1 and Patients 2 and 3 mentioned that this might put too much emphasis on their condition and that they do not want to be confronted with it too much:

> *Blood pressure measuring is a hot item, but it can make people feel unhappy as well. We should not exaggerate it, but it's definitely an important factor.* [GP 1]

Patients pointed out that measuring their blood pressure at home has benefits (6 quotes), but they do have insecurities about the self-measurement (15 quotes). Patients are insecure about how to perform a measurement and self-interpret the results, and they are anxious about the accuracy of the product:

> *Every time I measure, I wonder: do I measure technically good? Do I operate the product in the right way? Do I measure at the right moment?* [Patient 1]

Theme 4: New Touchpoints Between Patients and Doctors

To ensure the successful implementation of the eHealth product, new touchpoints between the different key actors were defined. Both the patients and the GPs want to keep scheduling regular

consulting moments, which can be with the GP or the GP assistant. From the analysis of all co-creation sessions, we concluded that patients prefer to measure their own blood pressure at home but would also like to have contact with their GP occasionally. All patients mentioned that they would prefer *data exchange* of the results through an app (39 quotes) or email and would prefer occasional feedback on the data overviews from a care professional, either the GP or GP assistant (29 quotes). The GPs are also in agreement that there should be personal contact between a GP and patient because:

> ...*you can use personal contact to motivate patients.* [GP 2]

The patients want to be able to view and analyze their own data to have a sense of control (13 quotes):

> *I like the feeling that I can look at my own results. That is important to me ... I want to keep an eye on my GP.* [Patient 1]

Concerning the frequency of contact moments, interestingly the patient with eHealth experience wanted to be checked more frequently and preferred having this done by the GP:

> ...*a GP would have probably responded to that, but the assistant did not.* [Patient 1]

The others agreed that the current frequency of GP consultations, every 6 or 12 months, is preferable (28 quotes), as they are not then constantly reminded about their blood pressure. There is a further division in whom patients want to share and discuss their data with. Patients 2 and 3 were satisfied with having a consultation once a year done by a GP assistant, under the supervision of the GP, who has greater expertise. The participants also wanted to keep the specialist in the loop to help with the interpretation of the blood pressure data concerning the risks of their particular health condition (9 quotes). Overall, we concluded that participants would prefer a Care Pathway that includes both self-measurement and professional care monitoring.

Integrated Care Pathway Design

The new design of the Care Pathway for at-risk patients as secondary preventive care is displayed in Figure 4. As in the current situation, the patient visits the GP, and their blood pressure is measured by the GP assistant. The design of the pathway differs in that it provides a new secondary preventive health care service:

- When the blood pressure is too high, the patient will receive a blood pressure monitor for at-home use and start to self-measure the blood pressure, generating an overview over a period of 3 months.
- The measured data will be saved by the device and automatically transferred to the health record system.
- The GP assistant receives an e-message if the patient's blood pressure is too high.
- Instantly the GP assistant analyzes the data and provides direct feedback to the patient.
- In some incidents, this feedback message might state that the data needs to be checked by the GP.
- For most incidents, the GP assistant is expected to do the monitoring and communicates with the patient.
- In stable situations, there is a (digital) contact moment every 3 months.
- Furthermore, the GP assistant will also communicate with the pharmacists on changes in medication doses.
- The blood and medication monitoring remain under the supervision of the GP, who will make the final decisions concerning medication and further diagnoses and treatments.
- Another aspect that is similar to the current process is that if the GP cannot control the blood pressure or if a patient has many other risk factors, the GP will refer the patient to a specialist (in most cases the internist). The specialist will examine the patient further.

Figure 4. Care pathway of blood pressure monitoring for preventive care.
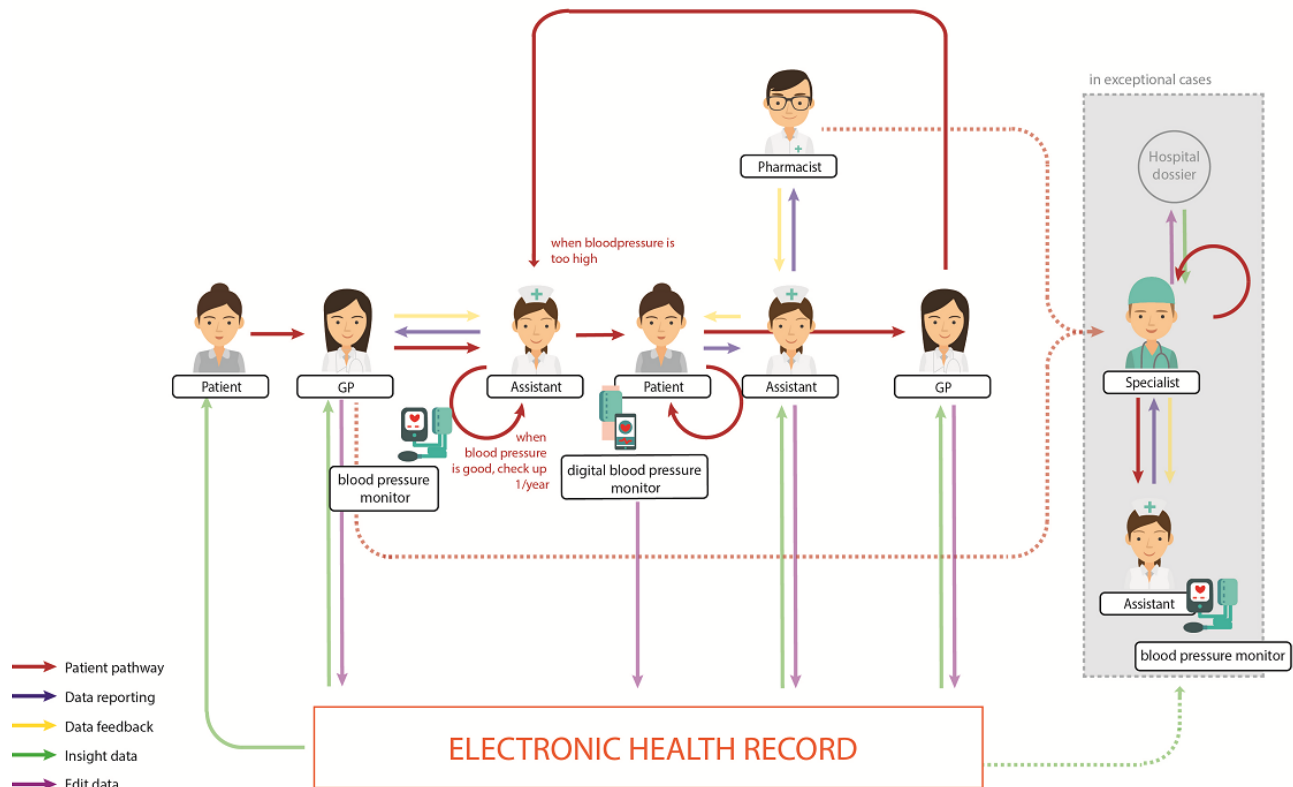
Discussion

Principal Results

This study developed a Care Pathway with a new secondary preventive care service that embeds at-home blood pressure measurement activities into the health care organization of diagnoses and treatment of at-risk patients. The study showed that in the current situation the GP measures the blood pressure during a consultation and prescribes medicines when the blood pressure is too high after a couple of measurements. This routine activity for at-risk patients is considered to belong to tertiary preventive care, where the care for patients is curative and reimbursed against higher costs. This Care Pathway design realizes a service shift to secondary preventive care, in which incidents and additional diseases can be tracked down in an earlier stage.

First, the routine activity is performed at the source, the patient is self-involved in executing the blood pressure measurements, leading to a feeling of being in control but also to a commonly acknowledged need to exchange and share the data with professional care providers. Second, this new touchpoint of data sharing and professional monitoring is performed by GP assistants with the additional benefit of earlier detection and reinforcement of lifestyle changes. Third, the data reporting is automated from the eHealth device into the health record system, integrating the patient as an implicit user of the health record system. A principal aspect of the design of a Care Pathway is the gathering of in-depth insights, not only on the current process but also on the future of health care services envisioned in co-creation sessions.

Comparison With Prior Work and Theoretical Implications

With this pathway design study, we showed that the activity of blood pressure monitoring can be shifted to the patient who can use an eHealth device at home under the precondition that the use is embedded into the Care Pathway, with a point of contact and safety service of incident monitoring at the organization of the GP practice.

These in-depth findings confirm the principal drive behind the patient-centered movement in which patients are actively involved and take greater responsibility for their own health [2]. This pathway design incorporates an eHealth blood pressure device that gives patients immediate feedback and insights into their health data, providing a direct impulse to improve their health outcomes. Patients can first try to manage or reduce their blood pressure themselves, instead of relying immediately on medication prescriptions.

Reasons to measure blood pressure at home have been researched in other studies [20,43]. An initial benefit of measurement at home has already been confirmed by previous research that found lower measurement outcomes of blood pressure self-measurement. Research showed that the elimination of the white-coat effect results in better and more accurate measurement data [20]. However, the opportunity to shift from tertiary to secondary preventive care and the implementation opportunities have not been found in previous studies. This is the first study that presents a Care Pathway for blood pressure monitoring and provides a detailed overview of the findings from this research concerning a new secondary preventive health service. In the current process, GPs are not able to filter their systems for at-risk patients and therefore

secondary preventive care can be started only after the patient has been in contact with the GP and high blood pressure is detected. With the use of the Care Pathway, more data about the patient's blood pressure is gathered and available. Both the patient and GP can review the measured data. The device has become part of the data model. It automatically transfers the data to the health record system, even when patients perform the measurement themselves. The Care Pathway design can help overcome implementation barriers [17]. The Care Pathway displays the touchpoints and role division between patients and health care professionals, which can help solve the lack of collaboration between health care parties.

The Care Pathway visualizes a clear role division and indicates that high involvement is needed to increase collaboration between the parties. Prior research found implementation barriers related to the shortage of funding [13]. To successfully implement blood pressure eHealth for patients at home, the financial responsibilities must be specified in detail. In support of this, the pathway design can now serve as a boundary object—an object with the capacity of translating and transferring between different viewpoints—to discuss the financial exchanges between the roles and organizations involved in the embedding of eHealth services. This next step in the embedding of eHealth services into the organization is also part of broader research efforts across types of eHealth devices and across countries and health care systems that focus on the implications and adaptations of the service and financial models. With the comparison of similar studies on the design of care pathways for eHealth devices at home, a next frontier in research is to develop comprehensive service models based on the data exchanges and transaction between the different parties in the networks.

Limitations and Implications for Further Research

Despite the strengths of generative sessions with multiple stakeholders, several limitations need to be discussed. The Care Pathway created in the current study is primarily based on the Dutch health care system. For implementation in other countries, additional design research is required. To repeat the method of inquiry and reuse the visual mapping toolkit to redesign clinical workflows of blood pressure self-measurement into care pathways in other countries in Europe, the United States and Asia will add to the body of knowledge on care pathway design and the generalizability of these designs.

In design research, and in particular in generative sessions, the data obtained are rich and mainly dependent on the context. The sample size of this in-depth qualitative design research was limited to 6 participant sites, offering enough robustness for the pathway design. In this qualitative research, we used words and drawings as the main data source. Both need to be interpreted

subjectively by the design researchers. The means of seeking objectivity that we used in this research were triangulation and multiple coding [38]. However, an important consideration in using the toolkit was to ensure that the facilitators remained objective and would not influence the "experts" while doing the mapping session. Furthermore, in our data analysis we came across differences between the patients in terms of the seriousness of their condition, their experiences of the role the disease has in their life, and their attitude towards monitoring. Some patients were more motivated to perform frequent measurements than others. Therefore, we suggest further specifying the at-risk patient group of the Care Pathway with additional quantitative research, such as a survey to investigate which patients benefit the most from the use of eHealth blood monitoring. A database query is also recommended.

A next avenue of research is also to organize pilot studies on to what extent lifestyle adaptations through monitoring can contribute to lowering a patient's blood pressure. Furthermore, the Care Pathway could benefit from additional research that weighs the costs versus benefits of blood pressure self-measurement with eHealth [24] and focuses on the effects of the Care Pathway on cost reductions and efficiency. A final implication for further research concerns the comparison of similar studies on the design of care pathways for eHealth devices at home in order to further theorize on the principles, concept, and frameworks that are useful for the embedding of eHealth interventions in health care organizations.

Practical Implications

For health professionals involved in eHealth innovation, this paper provides an example of a pathway model design that embeds eHealth technologies into an integrated service. The design method of generative sessions with the visual toolkit enabled the co-design of the example eHealth services. The pathway design method enabled the embedding of the eHealth devices into the service, providing the organization a person-centered perspective.

Conclusion

This research resulted in a Care Pathway of blood pressure monitoring for at-risk patients as secondary preventive care. The Care Pathway was designed to guide the implementation of eHealth devices for self-measurement of blood pressure. It showcases the pathway of at-risk patients and increases their involvement in managing their blood pressure. Furthermore, the Care Pathway leads to more accurate and reliable blood pressure data about patients, which could contribute to lower use of medicines and better insight into lifestyle influences on blood pressure. The Care Pathway serves as a basis for a new service that uses eHealth in future health care.

XSL•FO

RenderX

Conflicts of Interest

None declared.

References

1. McMullan M. Patients using the Internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
2. McColl-Kennedy JR, Snyder H, Elg M, Witell L, Helkkula A, Hogan SJ, et al. The changing role of the health care customer: review, synthesis and research agenda. Journal of Service Management 2017 Mar 20;28(1):2-33 [FREE Full text] [doi: 10.1108/JOSM-01-2016-0018]
3. Wildevuur SE, Simonse LWL. Information and communication technology-enabled person-centered care for the "big five" chronic conditions: scoping review. J Med Internet Res 2015 Mar 27;17(3):e77 [FREE Full text] [doi: 10.2196/jmir.3687] [Medline: 25831199]
4. European Commission. European Citizens: Digital Health Literacy. URL: http://ec.europa.eu/commfrontoffice/publicopinion/flash/fl_404_en.pdf [accessed 2019-03-13] [WebCite Cache ID 76r8DY9Yk]
5. Coulter A. Evidence based patient information. BMJ 1998 Jul 25;317(7153):225-226 [FREE Full text] [doi: 10.1136/bmj.317.7153.225]
6. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7 [FREE Full text] [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
7. Wolf A, Moore L, Lydahl D, Naldemirci O, Elam M, Britten N. The realities of partnership in person-centred care: a qualitative interview study with patients and professionals. BMJ Open 2017 Jul 17;7(7):e016491 [FREE Full text] [doi: 10.1136/bmjopen-2017-016491] [Medline: 28716793]
8. Irfan Khan A, Gill A, Cott C, Hans PK, Steele Gray C. mHealth Tools for the Self-Management of Patients With Multimorbidity in Primary Care Settings: Pilot Study to Explore User Experience. JMIR Mhealth Uhealth 2018 Aug 28;6(8):e171 [FREE Full text] [doi: 10.2196/mhealth.8593] [Medline: 30154073]
9. Epstein RM, Street RL. The values and value of patient-centered care. Ann Fam Med 2011;9(2):100-103 [FREE Full text] [doi: 10.1370/afm.1239] [Medline: 21403134]
10. Beard JR, Bloom DE. Towards a comprehensive public health response to population ageing. The Lancet 2015 Feb;385(9968):658-661 [FREE Full text] [doi: 10.1016/S0140-6736(14)61461-6]
11. World Health Organization. Global diffusion of eHealth: Making universal health coverage achievable. 2016. URL: https://www.who.int/goe/publications/global_diffusion/en/ [accessed 2019-03-13] [WebCite Cache ID 76r9BkqEd]
12. World Health Organization. A global brief on hypertension: silent killer, global public health crisis. 2013. URL: https://www.who.int/cardiovascular_diseases/publications/global_brief_hypertension/en/ [accessed 2019-03-13] [WebCite Cache ID 76r9LPsiO]
13. Bergmo TS. How to Measure Costs and Benefits of eHealth Interventions: An Overview of Methods and Frameworks. J Med Internet Res 2015 Nov 09;17(11):e254 [FREE Full text] [doi: 10.2196/jmir.4521] [Medline: 26552360]
14. Wilson C, Hargreaves T, Hauxwell-Baldwin R. Smart homes and their users: a systematic analysis and key challenges. Pers Ubiquit Comput 2014 Sep 4;19(2):463-476 [FREE Full text] [doi: 10.1007/s00779-014-0813-0]
15. Balta-Ozkan N, Davidson R, Bicket M, Whitmarsh L. Social barriers to the adoption of smart homes. Energy Policy 2013 Dec;63:363-374 [FREE Full text] [doi: 10.1016/j.enpol.2013.08.043]
16. Nieboer ME, van Hoof J, van Hout AM, Aarts S, Wouters EJM. Professional values, technology and future health care: The view of health care professionals in The Netherlands. Technology in Science 2014;39:10-17 [FREE Full text] [doi: 10.1016/j.techsoc.2014.05.00]
17. Ekblaw A, Azaria A, Halamka JD, Lippman A. MIT Media Lab. 2016. A Case Study for Blockchain in Healthcare:"MedRec" prototype for electronic health records and medical research data URL: https://www.healthit.gov/sites/default/files/5-56-onc_blockchainchallenge_mitwhitepaper.pdf [accessed 2019-03-13] [WebCite Cache ID 76rACnGOr]
18. van Meeuwen DPD, van Walt Meijer QJ, Simonse L. Care Models of eHealth Services: A Case Study on the Design of a Business Model for an Online Precare Service. JMIR Res Protoc 2015 Mar 24;4(1):e32 [FREE Full text] [doi: 10.2196/resprot.3501] [Medline: 25831094]
19. Crowley MJ, Bosworth HB, Coffman CJ, Lindquist JH, Neary AM, Harris AC, et al. Tailored Case Management for Diabetes and Hypertension (TEACH-DM) in a community population: study design and baseline sample characteristics. Contemp Clin Trials 2013 Sep;36(1):298-306 [FREE Full text] [doi: 10.1016/j.cct.2013.07.010] [Medline: 23916915]
20. O'Brien E, Asmar R, Beilin L, Imai Y, Mallion J, Mancia G, European Society of Hypertension Working Group on Blood Pressure Monitoring. European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement. J Hypertens 2003 May;21(5):821-848 [FREE Full text] [doi: 10.1097/01.hjh.0000059016.82022.ca] [Medline: 12714851]
21. Pickering TG. How Common Is White Coat Hypertension? JAMA 1988 Jan 08;259(2):225. [doi: 10.1001/jama.1988.03720020027031]

22. Rijksinstituut voor Volksgezondheid en Milieu. Wat is preventie?. 2015. URL: https://rivm.nl/Onderwerpen/K/ Kosteneffectiviteit_van_preventie/Over_kosteneffectiviteit_en_preventie/Opbouw_interventies_literatuurdatabase/ Wat_is_preventie

23. NIDDK. 2014. High Blood Pressure & Kidney Disease URL: https://www.niddk.nih.gov/health-information/kidney-disease/ high-blood-pressure [accessed 2019-03-13] [WebCite Cache ID 76rAq46fK]

24. Mileski M, Kruse CS, Catalani J, Haderer T. Adopting Telemedicine for the Self-Management of Hypertension: Systematic Review. JMIR Med Inform 2017 Oct 24;5(4):e41 [FREE Full text] [doi: 10.2196/medinform.6603] [Medline: 29066424]

25. Nierstichting. Feiten en cijfers. 2018. URL: https://www.nierstichting.nl/leven-met-een-nierziekte/feiten-en-cijfers/ [accessed 2019-03-13] [WebCite Cache ID 76rB5QLX7]

26. Bettencourt-Silva JH, Clark J, Cooper CS, Mills R, Rayward-Smith VJ, de la Iglesia B. Building Data-Driven Pathways From Routinely Collected Hospital Data: A Case Study on Prostate Cancer. JMIR Med Inform 2015 Jul 10;3(3):e26 [FREE Full text] [doi: 10.2196/medinform.4221] [Medline: 26162314]

27. Oosterholt R, Simonse L, Boess S, Vehmeijer SBW. Designing a Care Pathway Model - A Case Study of the Outpatient Total Hip Arthroplasty Care Pathway. Int J Integr Care 2017 Mar 09;17(1):2 [FREE Full text] [doi: 10.5334/ijic.2429] [Medline: 29042844]

28. Vanhaecht K, Bollmann M, Bower K, Gallagher C, Gardini A, Guezo J, et al. Prevalence and use of clinical pathways in 23 countries – an international survey by the European Pathway Association. Journal of Integrated Care Pathways 2018 May 30;10(1):28-34 [FREE Full text] [doi: 10.1177/205343540601000106]

29. Cringles MC. Developing an integrated care pathway to manage cancer pain across primary, secondary and tertiary care. Int J Palliat Nurs 2002 May;8(5):247-255 [FREE Full text] [doi: 10.12968/ijpn.2002.8.5.10372] [Medline: 12048508]

30. Hall J. A qualitative survey of staff responses to an integrated care pathway pilot study in a mental healthcare setting. NT Research 2016 Aug 18;6(3):696-705 [FREE Full text] [doi: 10.1177/136140960100600309]

31. Rees G, Huby G, McDade L, McKechnie L. Joint working in community mental health teams: implementation of an integrated care pathway. Health Soc Care Community 2004 Nov;12(6):527-536 [FREE Full text] [doi: 10.1111/j.1365-2524.2004.00523.x] [Medline: 15717899]

32. Patton MQ. Qualitative Research & Evaluation Methods. New York, NY: Sage Publications, Inc; 2002.

33. Suter WN. Qualitative data, analysis, and design. In: Introduction to Educational Research: A Critical Thinking Approach, Second Edition. Thousand Oaks, CA: Sage Publications; 2012:366-367.

34. Visser FS, Stappers PJ, van der Lugt R, Sanders EB. Contextmapping: experiences from practice. CoDesign 2005 Apr;1(2):119-149 [FREE Full text] [doi: 10.1080/15710880500135987]

35. Creswell JW, Plano Clark VL. Designing And Conducting Mixed Methods Research. Thousand Oaks, CA: Sage Publications, Inc; 2011.

36. TU Delft. Human Research Ethics. URL: https://www.tudelft.nl/over-tu-delft/strategie/strategiedocumenten-tu-delft/ integriteitsbeleid/human-research-ethics/ [accessed 2019-03-13] [WebCite Cache ID 76rCIGqTZ]

37. Sanders L, Stappers JP. Convivial Toolbox: Generative Research for the Front End of Design. Amsterdam, Netherlands: BIS Publishers; 2012.

38. Patton MQ. Qualitative interviewing. In: Qualitative research & evaluation methods (3rd ed.). Thousand Oaks, CA: Sage Publications; 2002:339-418.

39. Maxwell JA. Understanding and Validity in Qualitative Research. Harvard Educational Review 1992 Sep;62(3):279-301 [FREE Full text] [doi: 10.17763/haer.62.3.8323320856251826]

40. Denzin N. The Research Act: A Theoretical Introduction to Sociological Methods. New York, NY: McGraw-Hill; 1978.

41. Saldaña J. First cycle coding methods. In: The coding manual for qualitative researchers. London, UK: Sage Publications; 2012.

42. Miles MB, Huberman AM, Saldaña J. Research Design and Management. In: Qualitative Data Analysis: A Methods Sourcebook. Thousand Oaks, CA: Sage Publications, Inc; 2013.

43. McKay DW, Godwin M, Chockalingam A. Practical advice for home blood pressure measurement. Can J Cardiol 2007 May 15;23(7):577-580 [FREE Full text] [Medline: 17534466]

Abbreviations

eHealth: electronic health
GP: general practitioner
HREC: Human Research Ethics Committee
ICT: information and communication technology

Original Paper

Mobile Health for Central Sleep Apnea Screening Among Patients With Stable Heart Failure: Single-Cohort, Open, Prospective Trial

Roderick Willem Treskes[1], MSc, PhD; Arie C Maan[1], MSc, PhD; Harriette Florence Verwey[1], MD, PhD; Robert Schot[2], BEng; Saskia Lambertha Maria Anna Beeres[1], MD, PhD; Laurens F Tops[1], MD, PhD; Enno Tjeerd Van Der Velde[1], MSc, PhD; Martin Jan Schalij[1], MD, PhD; Annelies Margaretha Slats[2], MD, PhD

[1]Department of Cardiology, Leiden University Medical Center, Leiden, Netherlands
[2]Department of Pulmonology, Leiden University Medical Center, Leiden, Netherlands

Corresponding Author:
Roderick Willem Treskes, MSc, PhD
Department of Cardiology
Leiden University Medical Center
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 0715262020
Fax: 31 715266809
Email: r.w.treskes@lumc.nl

Abstract

Background: Polysomnography is the gold standard for detection of central sleep apnea in patients with stable heart failure. However, this procedure is costly, time consuming, and a burden to the patient and therefore unsuitable as a screening method. An electronic health (eHealth) app to measure overnight oximetry may be an acceptable screening alternative, as it can be automatically analyzed and is less burdensome to patients.

Objective: This study aimed to assess whether overnight pulse oximetry using a smartphone-compatible oximeter can be used to detect central sleep apnea in a population with stable heart failure.

Methods: A total of 26 patients with stable heart failure underwent one night of both a polygraph examination and overnight saturation using a smartphone-compatible oximeter. The primary endpoint was agreement between the oxygen desaturation index (ODI) above or below 15 on the smartphone-compatible oximeter and the diagnosis of the polygraph.

Results: The median age of patients was 66.4 (interquartile range, 62-71) years and 92% were men. The median body mass index was 27.1 (interquartile range, 24.4-30.8) kg/m^2. Two patients were excluded due to incomplete data, and two other patients were excluded because they could not use a smartphone. Seven patients had central sleep apnea, and 6 patients had obstructive sleep apnea. Of the 7 (of 22, 32%) patients with central sleep apnea that were included in the analysis, 3 (13%) had an ODI≥15. Of all patients without central sleep apnea, 8 (36%) had an ODI<15. The McNemar test yielded a *P* value of .55.

Conclusions: Oxygen desaturation measured by this smartphone-compatible oximeter is a weak predictor of central sleep apnea in patients with stable heart failure.

KEYWORDS

Introduction

Central sleep apnea is characterized by sleep-disordered breathing associated with diminished or absent respiratory effort. It is often accompanied by symptoms of tiredness, excessive daytime sleepiness, and frequent nocturnal awakening [1,2].

Central sleep apnea and Cheyne-Stokes respiratory breathing are common in patients with congestive heart failure, with a reported prevalence of 30%-50% [3]. Moreover, central sleep apnea in chronic heart failure is associated with increased mortality and reduced left ventricular function [4]. In addition, treatment of central sleep apnea with continuous positive airway pressure in cases of chronic heart failure has shown to improve

left ventricular function in patients who respond to treatment [5].

Polysomnography is the gold standard for the diagnosis of central sleep apnea. However, it is a burden to the patient, as it disrupts sleep. Furthermore, the process is time consuming for technicians, as one polysomnography examination takes 2 hours to fully evaluate. The polygraph examination is an easier way of evaluating sleep-disordered breathing than full polysomnography, but the former also still disrupts normal sleep for a patient with an already reduced quality of life. Other screening methods to reduce the number of polygraphs may therefore be preferred.

Developing new screening methods, including electronic health (eHealth) apps, questionnaires, and wireless overnight pulse oximetry for patients with congestive heart failure might optimize the number of patients screened for central sleep apnea. Furthermore, it may be more patient friendly in a group of patients with an already diminished quality of life.

One possible screening method is the use of eHealth apps. Recent developments in the eHealth industry resulted in a variety of eHealth apps that claim that they can detect sleep-disordered breathing. However, most of these apps are not clinically validated. One example of an app that, according to the manufacturer, provides accurate saturation measurements is the iSpO$_2$ app (Masimo Corporation, Irvine, California) [6,7]. By using the app and an oximeter, a user can record saturation, heart rate, and pulse index. Digital storage of data allows for rapid transmission and analysis, minimizing the involvement of technicians. Previous research has suggested that overnight oximetry can be used to detect obstructive sleep apnea in various patient populations [8]. However, overnight oximetry with an eHealth device has not been evaluated as a screening method for patients with stable heart failure. Therefore, this study aimed to evaluate whether overnight pulse oximetry can be used to identify patients with central sleep apnea in a population with congestive heart failure by using a validated mobile health app.

Methods

Patient Population

Patients with stable heart failure who visited the outpatient clinic of the Department of Cardiology at the Leiden University Medical Center were eligible for study participation if they met all inclusion and exclusion criteria. The inclusion and exclusion criteria are listed in Textbox 1. Briefly, patients who had stable heart failure according to the European Society of Cardiology guidelines [9], no history of obstructive sleep apnea or central sleep apnea, no history of ischemic or hemorrhagic stroke, and a life expectancy of more than 12 weeks (as per the physician's discretion) were eligible.

Study Design and Procedure

The study was a single-cohort, nonrandomized, open, prospective trial. Patients with stable heart failure were asked to participate by a treating cardiologist at a regularly scheduled heart failure outpatient clinic visit. Patients received information from a project-dedicated health care professional. If patients were willing to participate, they visited the Department of Pulmonology within 1.5 months of the outpatient clinic visit. At day 1, a project-dedicated health care professional with ample training performed the polygraph. Each patient was given a smartphone and smartphone-compatible oximeter and received oral and written instructions on their use. Patients were instructed to attach the smartphone-compatible oximeter contralateral to the hand where the polygraph was attached. During the first night, patients slept with both the polygraph and the smartphone-compatible oximeter attached. After one night, patients returned the polygraph to the hospital. On the second, third, and fourth nights, patients slept with only the smartphone-compatible oximeter attached. After the fourth night, patients returned the smartphone-compatible oximeter to the hospital. A flowchart of these events is presented in Figure 1.
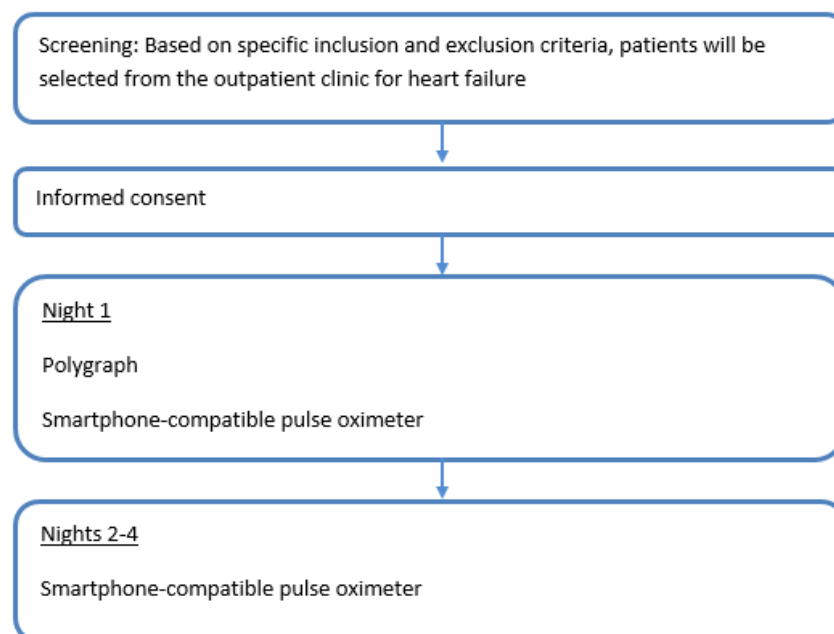
Devices

The polygraph equipment (Cidelec, Angers, France) consisted of a nasal cannula, a suprasternal sensor, thoracic and abdominal gauges, a finger pulse oximeter, a light sensor, body position sensor, and an activity sensor. The pulse oximeter has a sampling rate of 8 Hz. Both the smartphone (iPhone 5s; Apple, Cupertino, CA) and the smartphone-compatible pulse oximeter (Masimo) were provided by the hospital for the duration of the study. The smartphone-compatible pulse oximeter is worn at the fingertip and is connected with the smartphone via a wire. The pulse oximeter has a sampling rate of 1 Hz.

Devices used in this study were battery powered, electrically safe, and approved by the hospital's Instrumentation Department.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Chronic heart failure according to the European Society of Cardiology guidelines [9]

- Age≥18 years

Exclusion criteria:

- History of obstructive sleep apnea syndrome

- History of central sleep apnea syndrome

- History of ischemic stroke

- History of hemorrhagic stroke

- History of chronic obstructive pulmonary disease

- Evidence of fluid retention at the time of study inclusion

- History of surgery under general anesthesia ≤3 months before study inclusion

- Intravenous injection of diuretics ≤1 month before study inclusion

- Unwilling to sign the informed consent form

- Life expectancy ≤12 weeks as per the physician's discretion

- History of left ventricular assist device implantation

- Use of oxygen on a daily basis

- Pregnancy

Figure 1. Flowchart of study procedures.

Data Analysis

Central sleep apnea and obstructive sleep apnea were diagnosed on the basis of the polygraph examination results in accordance with the America Association of Sleep Medicine guidelines [10]. Patients were diagnosed with sleep apnea if the polygraph showed an apnea-hypopnea index of ≥15 per hour. Sleep apnea was subsequently classified as central or obstructive sleep apnea. A patient was diagnosed with central sleep apnea if ≥50% of all apnea and hypopnea events were classified as "central." A patient was diagnosed with obstructive sleep apnea if <50% of

all apnea and hypopnea events were classified as "central." Definitions for apnea and hypopnea events and their subdivision as central or obstructive were derived from the America Association of Sleep Medicine manual for the scoring of sleep and associated events [10]. The oxygen desaturation index (ODI) was defined as the average number of dips in saturation per hour. A cut-off value of 15 was chosen for the ODI. A dip was defined as a ≥3% decrease in saturation that lasted ≥10 seconds from the baseline saturation. The baseline saturation was determined in the hospital right after the polygraph was attached

to the patient. The polygraphs were reviewed by a senior pulmonary physician with ample training who was blinded to the results of the oximeter-compatible app.

The oximeter-compatible app (Masimo) generated a comma-separated value (CSV) file, which was imported into a dedicated Matlab script (The MathWorks, Natick, MA), from which the average pulse oximeter saturation (SpO_2), lowest SpO_2, total percentage of time spent with a saturation <90%, and ODI were calculated. The ODI was again defined as the average number of dips in saturation per hour. A dip was defined as a ≥3% decrease in saturation that lasted ≥10 seconds from the average saturation over the 11th minute of measurement [10]. The smartphone-compatible oximeter data were analyzed by a project-dedicated professional with ample training, who was blinded to the results of the polygraph.

Endpoints

The primary endpoint is an agreement between ODI of the smartphone-compatible oximeter and the diagnosis of the polygraph, expressed as four numbers (the number of patients who have central sleep apnea, as diagnosed by polygraph, and ≥15 dips/hour on the smartphone-compatible oximeter; the number of patients who have central sleep apnea and ≤15 dips/hour; the number of patients who do not have central sleep apnea and ≥15 dips/hour; the number of patients who do not have central sleep apnea and ≤15 dips/hour) in a 2 2 table.

The secondary endpoints include (1) the percentage of cases of sleep apnea (either of obstructive or central etiology) detected in the study population by the polygraph; (2) the percentage of cases of central sleep apnea detected in the study population by the polygraph; (3) agreement between the ODI, measured by the polygraph, and sleep apnea (either obstructive or central etiology) in the study population; (4) agreement between the ODI, measured by the polygraph, and central sleep apnea in the study population; (5) median difference in ODI, lowest saturation, and average saturation between the polygraph and mobile pulse oximeter; (6) sensitivity and specificity of pulse oximetry to detect central sleep apnea by saturation dips>15/hour; and (7) the percentage of patients able to use the eHealth device as instructed.

Statistical Analysis

R (R foundation for statistical computing, Vienna, Austria) was used to perform a power calculation for the McNemar test. An alpha level of .05 and a beta level of 0.20 were chosen. Based on unpublished research by our study group, we estimated the ratio of p01/p10 (p01, false positives; p10, false negatives) to be 12, and the sum of p10 and p01 to be 0.39. This yielded a sample size of 26 patients.

SPSS version 22.0 (IBM Corp, Armonk, NY) was used for statistical analysis. Continuous variables are expressed as median with interquartile range (IQR) from the 25th to the 75th percentile.

Significance of the primary endpoint was calculated using the McNemar test. A P value≤.05 was considered statistically significant. A Bland-Altman plot was drafted to assess the short-term reproducibility of the ODI, with the ODI of the first night depicted on the X-axis and the difference in the ODI between the first and second night depicted on the Y-axis.

Ethical Approval

This study was conducted in accordance to the principles of the Declaration of Helsinki (version 10, October 2013) [11] as per the Dutch Medical Research Involving Human Subjects Act [12] and Good Clinical Practice [13]. The study was approved by the hospital's Medical Ethics Committee (P15.211). All subjects provided written informed consent before inclusion in the study. Devices used in this study were all battery powered and electrically safe. All devices were approved by the hospital's Instrumentation Department. All devices used in this study were purchased from manufacturers. No manufacturer had any role in or influence on the design or conduct of the study, data analysis, writing of the manuscript, or the decision to submit for publication. No financial support for this study was received from any manufacturer.

Results

Patient Population

A total of 26 patients were included in the study. The median age was 66.4 (IQR: 62-71) years, and 92% were men. The median body mass index was 27.1 (IQR: 24.4-30.8) kg/m^2. All patients had New York Heart Association class I (15.4%) or class II (84.6%) heart failure, and 61.5% had an ischemic cardiomyopathy. The median left ventricular ejection fraction was 34% (IQR: 24%-45%), median probrain natriuretic peptide level was 748 (IQR: 244.6-1479) ng/L, and median neck circumference was 41 (IQR: 38-44) cm. The population characteristics are summarized in Table 1.

Polygraph

A total of 26 polygraph examinations were performed. One polygraph examination was of insufficient diagnostic quality and one polygraph examination was too short to establish a diagnosis. Both patients were not willing to undergo a second polygraph examination. Of the 24 patients that underwent a polygraph examination of diagnostic quality, 14 (58%) had sleep apnea (of either etiology); in addition, 8 of the 24 (33%) were diagnosed with central sleep apnea and 6 (25%) were diagnosed with obstructive sleep apnea (secondary endpoints 1 and 2 [see Endpoints section]). In 10 of the 24 (41%) cases, no sleep apnea was detected. The median sleep duration was 6.5 (IQR: 5.4-7.4) hours, median apnea-hypopnea index was 17 (IQR: 6.5-27.8), median ODI was 16 (IQR: 5.5-28), median number of hypopneas per night was 62 (IQR: 30.8-79.8), and median number of dips was 92.5 (IQR: 33.3-156).

XSL•FO
RenderX

Table 1. Baseline characteristics of the study population (N=26).

Characteristics	Value
Age (years), IQR[a]	66.4 (62.2-70.6)
Male gender, n (%)	24 (92.3)
Body mass index (kg/m^2), median (IQR)	27.1 (24.4-30.8)
NYHA[a] class, n (%)	
I	4 (15.4)
II	22 (84.6)
Ischemic cardiomyopathy, n (%)	16 (61.5)
Left ventricular ejection fraction, median (IQR)	34 (23.5-45)
Pro-brain natriuretic peptide level, median (IQR)	748 (244.6-1479)
Neck circumference (cm), median (IQR)	41 (35-49)

[a]IQR: interquartile range.

[b]NYHA: New York Heart Association.

Overnight Oximetry

All 26 participants transferred at least one CSV file containing the overnight saturation measured by the smartphone-compatible pulse oximeter. Of the 4 patients who did not transfer a CSV file of their first night (the night they also underwent the polygraph), 2 patients could not be diagnosed because the polygraph was of insufficient quality (as described above). The other 2 patients forgot to attach the smartphone-compatible pulse oximeter on their first night. Therefore, 22 patients were included in the analysis of the primary endpoint and secondary endpoints 3, 4, 5 and 6 (see Endpoints section; Figure 2). Of the 26 patients who participated, 13 (50%) were able to transfer CSV files of four consecutive nights. A total of 9 (35%) patients transferred CSV files of three nights, 1 (4%) transferred CSV files of two nights, and 3 (12%) transferred CSV files of only one night.

Of all files transferred, the median saturation was 95.7 (IQR: 94.5-96.7). The median lowest saturation was 87 (IQR: 82-90), and the median ODI was 10.1 (IQR: 2.9-20.3). The total number of dips was 75.5 (21-144.8), and total sleep time was 8.1 (6.7-9.4) hours.

Primary Endpoint

Of the 7 (of 22, 32%) patients with central sleep apnea who were included in the analysis, 3 (14%) had an ODI≥15. The other 4 (18%) patients had an ODI<15. Of all 15 patients without

central sleep apnea who were included in the analysis, 8 (36%) had an ODI<15 (Table 2). The McNemar test yielded a *P* value of .55.

Secondary Endpoints

Of all 13 patients with sleep apnea, 6 had an ODI≥15, measured by Masimo. Of all 9 patients without sleep apnea, 5 had an ODI<15 (Table 3).

Of the 7 patients with central sleep apnea, 6 had an ODI ≥15 (measured by the polygraph). Of the patients without central sleep apnea, 9 had an ODI<15 (Table 4). Of all 13 patients with sleep apnea (of either etiology), 12 had an ODI≥15 (measured by the polygraph). All 9 patients without sleep apnea had an ODI<15 (Table 5).

The sensitivity of the ODI for the detection of central sleep apnea is 43%, and the specificity is 53%. The positive predictive value is 30%, and the negative predictive value is 67%. The sensitivity of the polygraph is 86%, and the specificity is 53%.

Difference Between the Polygraph and Mobile Pulse Oximeter

The median difference in ODI was 2.1 indices higher than that measured with the polygraph (IQR: –4.3 to 14.3). The smartphone-compatible oximeter yielded a median saturation of 3.1 (IQR: 2.6-4.1) percentage points higher than the that measured by the polygraph. There was no difference in the median lowest saturation measured by both devices (both 83%).

Figure 2. Flowchart of the results of the primary outcome. PG: polygraph; ODI: oxygen desaturation index.
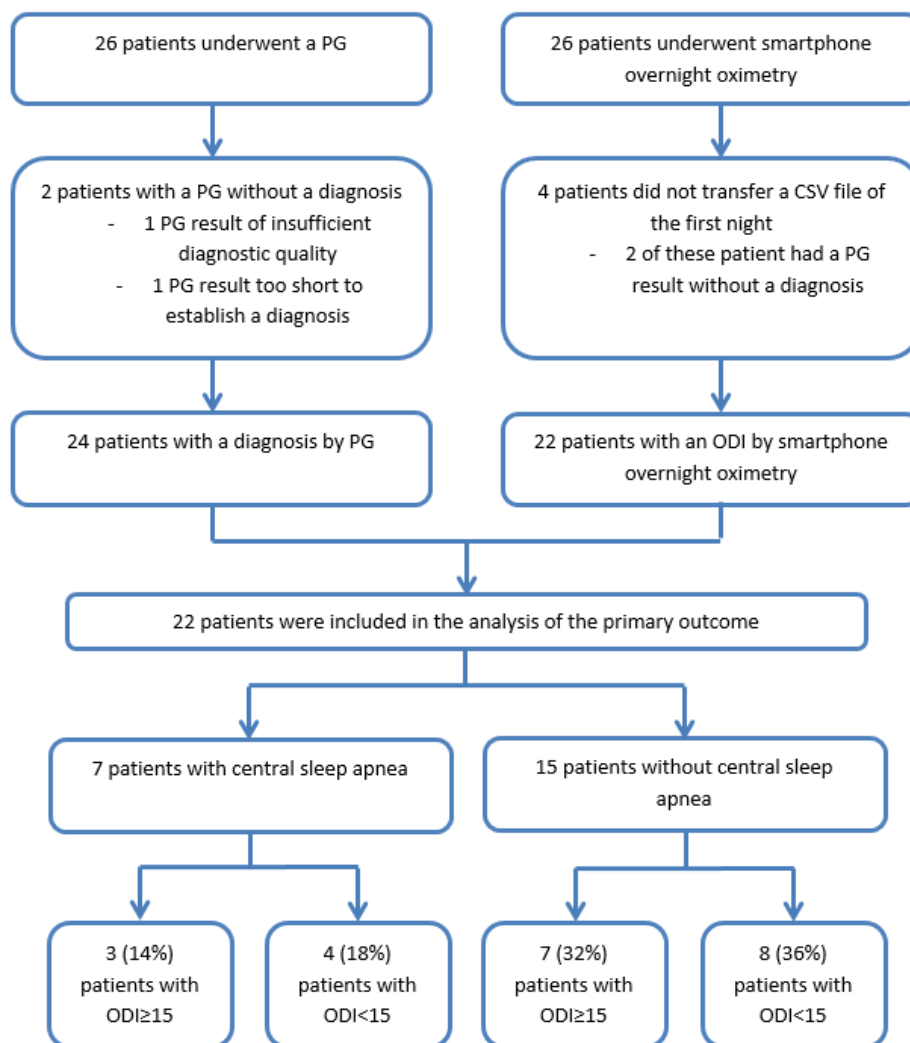
Table 2. Number of patients with central sleep apnea (yes/no) and an ODI of ≥15 or <15 (as measured by the smartphone-compatible oximeter).

Oxygen desaturation index measured by mobile pulse oximeter	Central sleep apnea present, n (%)	Central sleep apnea absent, n (%)	Total, n (%)
≥15	3 (14)	7 (32)	10 (46)
<15	4 (18)	8 (36)	12 (54)
Total	7 (32)	15 (68)	22 (100)

Table 3. Number of patients with sleep apnea (yes/no) and an ODI of ≥15 or <15 (as measured by the smartphone-compatible pulse oximeter).

Oxygen desaturation index measured by smart-phone-compatible oximeter	Sleep apnea present, n (%)	Sleep apnea absent, n (%)	Total, n (%)
≥15	6 (27)	4 (18)	10 (45)
<15	7 (32)	5 (23)	12 (55)
Total	13 (59)	9 (41)	22 (100)

Table 4. Number of patients with central sleep apnea (yes/no) and an ODI of ≥15 or <15 (as measured by the polygraph).

Oxygen desaturation index measured by polygraph	Central sleep apnea present, n (%)	Central sleep apnea absent, n (%)	Total, n (%)
≥15	6 (27)	6 (27)	12 (54)
<15	1 (5)	9 (41)	10 (46)
Total	7 (32)	15 (68)	22 (100)

Table 5. Number of patients with sleep apnea (yes/no) and an ODI of ≥15 or <15 (as measured by the polygraph).

Oxygen desaturation index measured by polygraph	Sleep apnea present, n (%)	Sleep apnea absent, n (%)	Total, n (%)
≥15	12 (54)	0 (0)	12 (54)
<15	1 (5)	9 (41)	10 (46)
Total	13 (59)	9 (41)	22 (100)

Discussion

Principal Findings

This study investigated the use of a smartphone-compatible oximeter to measure the ODI for the detection of central sleep apnea in patients with stable heart failure. Oxygen desaturation, when measured by an eHealth device, appeared to be a weak predictor of central sleep apnea in patients with stable heart failure. In addition, in this elderly group of patients, the correct use of this eHealth device was only achieved by 50% of patients. On the other hand, ODI measured by the polygraph might be a good predictor of sleep apnea of any etiology in patients with stable heart failure. We found that 58% of participating patients had sleep apnea. Of all patients, 33% had central sleep apnea and 25% had obstructive sleep apnea. These percentages are lower than the prevalence reported by Oldenburg et al [14]: In a screening study of 700 patients, they showed that 70% of patients had sleep apnea, of which 40% had central sleep apnea and 30% had obstructive sleep apnea. This difference may be explained (at least partly) by the relatively small sample size, but possibly also by the differences in severity of heart failure in these patients.

In our study, a median number of 92.5 dips was observed. This number was significantly higher than that in a previous study [15] of 12 patients with heart failure, which reported 4 dips per patient per night. This difference may largely be due to the difference in definition of a "dip." Davies et al [15] defined a dip as "a fall of >4% in oxygen saturation from a stable baseline that lasted >30 seconds," while in the current study, a dip was defined as a ≥3% decrease lasting 10 seconds. This criterion was necessary to define dips equally between the polygraph software and our smartphone-compatible oximeter.

Oxygen Desaturation Index as a Potential Screening Tool for Central Sleep Apnea

This study showed that the ODI, measured by either the polygraph or smartphone-compatible oximeter, correlates poorly with the diagnosis of central sleep apnea. The McNemar test yielded a nonsignificant *P* value of .55. However, there was a good correlation between the ODI measured by the polygraph and the diagnosis of sleep apnea (of either etiology). We acknowledge that the study was not powered on this outcome. Furthermore, hypopneas in polygraphs are scored based on desaturation events. Therefore, the diagnosis of sleep apnea is partially dependent on the ODI. However, the strong outcome of 0 false positives and 1 false negative indicates that the ODI might indeed be a good screening method and should be investigated in further research.

Implications for Clinical Practice

Our study found that 58% of patients with stable heart failure had either form of sleep apnea (central or obstructive etiology). Both obstructive and central sleep apnea are associated with higher mortality and lower quality of life in patients with stable heart failure [4]. Therefore, early diagnosis is of paramount importance. However, screening for obstructive sleep apnea or central sleep apnea is not recommended by current guidelines, but with such a high prevalence, routine screening of patients should be considered. Perhaps, screening should not focus on the distinction between obstructive sleep apnea and central sleep apnea, as both have clinical implications [9]. Since ODI correlates well with sleep apnea of any etiology, research on easy-to-use overnight oximetry eHealth devices for sleep apnea screening among patients with stable heart failure is necessary.

Electronic Health Use in a Population With Heart Failure

In this study, patients were asked to attach, record, and email the overnight saturations themselves. Instructions about the use of the smartphone, the Masimo patch, and the emailing of the CSV files were given after the polygraph was attached to the patient. However, of all patients, 13 were unable to transfer four CSV files. These results should be seen as hypothesis generating, but also indicate that when conducting a study in an older and vulnerable population, the eHealth system should be tailored to the patient population. Furthermore, time spent in patient education of the eHealth system should not be underestimated.

Differences in Saturation Measured by the Polygraph and the Mobile Pulse Oximeter

Our results showed some significant differences in the predictive value of the ODI for both sleep apnea of any etiology and central sleep apnea, median ODI, and median average saturation between the polygraph and the smartphone-compatible pulse oximeter. There are several explanations for this phenomenon. First, it is unclear how motion artefacts influenced our results. Motion during sleep and movements of the fingers might result in different results from the oximeter of the polygraph, which has been designed specifically for overnight saturation measurement. Second, patients attached the smartphone-compatible oximeter themselves at home. Although instructions were given in the hospital, it is uncertain whether patients attached the device correctly. Improper placement usually gives no signal and therefore no saturation in the CSV file. On the other hand, slight improper placement might result in improper values in the CSV file.

Limitations

This study has some limitations that have affected its results. Unfortunately, in two patients, it was not possible to obtain a

diagnosis from the polygraph. These two patients were not willing to undergo a second polygraph. However, given the numbers of the primary endpoint and a relatively high *P* value of .55, it is unlikely that two extra patients would have changed the data significantly. Furthermore, some patients could not deal with the smartphone technology given, despite ample instructions. As a consequence, 13 patients were unable to record their overnight saturation for four consecutive nights. Lastly, we did not perform polygraph examinations in healthy volunteers. Therefore, a comparison of overnight oximetry with aged-matched healthy volunteers is lacking, although this has been previously reported in the literature [16].

Conclusions

Oxygen desaturation, when measured by the eHealth oximeter tested in this study is a weak predictor of central sleep apnea in patients with stable heart failure. The ODI, when measured by a validated device, might be a good predictor of sleep apnea of any etiology in patients with stable heart failure. This study also corroborated the high prevalence of sleep apnea in patients with stable heart failure. Therefore, more research on screening for sleep apnea in patients with stable heart failure is warranted, which might be possible by using validated overnight oximetry, but must be easy to perform in this type of elderly patient group.

Conflicts of Interest

None declared.

References

1. Block A, Border W, Bruce B, Christopher K, Drake R, Jangid A. Dorland's Medical Dictionary. Philadelphia, Pennsylvania: Saunders; 2011.
2. Aurora RN, Chowdhuri S, Ramar K, Bista SR, Casey KR, Lamm CI, et al. The treatment of central sleep apnea syndromes in adults: practice parameters with an evidence-based literature review and meta-analyses. Sleep 2012 Jan 01;35(1):17-40 [FREE Full text] [doi: 10.5665/sleep.1580] [Medline: 22215916]
3. Costanzo MR, Khayat R, Ponikowski P, Augostini R, Stellbrink C, Mianulli M, et al. Mechanisms and clinical consequences of untreated central sleep apnea in heart failure. J Am Coll Cardiol 2015 Jan 06;65(1):72-84 [FREE Full text] [doi: 10.1016/j.jacc.2014.10.025] [Medline: 25572513]
4. Javaheri S, Shukla R, Zeigler H, Wexler L. Central sleep apnea, right ventricular dysfunction, and low diastolic blood pressure are predictors of mortality in systolic heart failure. J Am Coll Cardiol 2007 May 22;49(20):2028-2034 [FREE Full text] [doi: 10.1016/j.jacc.2007.01.084] [Medline: 17512359]
5. Arzt M, Floras J, Logan A, Kimoff R, Series F, Morrison D, CANPAP Investigators. Suppression of central sleep apnea by continuous positive airway pressure and transplant-free survival in heart failure: a post hoc analysis of the Canadian Continuous Positive Airway Pressure for Patients with Central Sleep Apnea and Heart Failure Trial (CANPAP). Circulation 2007 Jun 26;115(25):3173-3180. [doi: 10.1161/CIRCULATIONAHA.106.683482] [Medline: 17562959]
6. iSpO2 pulse oximeter. URL: http://masimopersonalhealth.com/ispo2/ [accessed 2018-01-18] [WebCite Cache ID 72h926AZH]
7. Huizing MJ, Villamor-Martínez E, Chavagne IA, Vanagt WY, Spaanderman MAE, Villamor E. Reliability and Validity of a Smartphone-Paired Pulse Oximeter for Screening of Critical Congenital Heart Defects in Newborns. Neonatology 2017;112(4):324-329. [doi: 10.1159/000477294] [Medline: 28768288]
8. Yalamanchali S, Farajian V, Hamilton C, Pott TR, Samuelson CG, Friedman M. Diagnosis of obstructive sleep apnea by peripheral arterial tonometry: meta-analysis. JAMA Otolaryngol Head Neck Surg 2013 Dec;139(12):1343-1350. [doi: 10.1001/jamaoto.2013.5338] [Medline: 24158564]
9. Ponikowski P, Voors A, Anker S, Bueno H, Cleland J, Coats A. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. European Heart Journal 2016;37(27):2129-2200.
10. Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, American Academy of Sleep Medicine. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med 2012 Oct 15;8(5):597-619 [FREE Full text] [doi: 10.5664/jcsm.2172] [Medline: 23066376]
11. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
12. Dutch Government. 1998. Wet medisch-wetenschappelijk onderzoek met mensen URL: https://wetten.overheid.nl/BWBR0009408/2017-03-01 [accessed 2019-02-07] [WebCite Cache ID 761EScg2Q]
13. Central Committee on Research Involving Human Subjects. 1993. Guidelines for Good Clinical Practice URL: http://www.ccmo.nl/attachments/files/ich-gcp-richtsnoer-gcp-vertaald-mrt-2003.pdf [accessed 2017-05-10]
14. Oldenburg O, Lamp B, Faber L, Teschler H, Horstkotte D, Töpfer V. Sleep-disordered breathing in patients with symptomatic heart failure: a contemporary study of prevalence in and characteristics of 700 patients. Eur J Heart Fail 2007 Mar;9(3):251-257 [FREE Full text] [doi: 10.1016/j.ejheart.2006.08.003] [Medline: 17027333]

15. Davies S, John L, Wedzicha J, Lipkin D. Overnight studies in severe chronic left heart failure: arrhythmias and oxygen desaturation. Br Heart J 1991 Feb;65(2):77-83 [FREE Full text] [Medline: 1907836]

16. Gries R, Brooks L. Normal oxyhemoglobin saturation during sleep. How low does it go? Chest 1996 Dec;110(6):1489-1492. [Medline: 8989066]

Abbreviations

CSV: comma-separated value
eHealth: electronic health
IQR: interquartile range
NYHA: New York Heart Association
ODI: oxygen desaturation index
SpO$_2$: pulse oximeter saturation

XSL•FO
RenderX

XSL•FO

RenderX