Original Papers

Cloud-Based Machine Learning Platform to Predict Clinical Outcomes at Home for Patients With Cardiovascular Conditions Discharged From Hospital: Clinical Trial (e45130)
Phillip Yang, Alokkumar Jha, William Xu, Zitao Song, Patrick Jamp, Jeffrey Teuteberg ................................................................. 2

Formative Perceptions of a Digital Pill System to Measure Adherence to Heart Failure Pharmacotherapy: Mixed Methods Study (e48971)
Peter Chai, Jenson Kaithamattam, Michelle Chung, Jeremiah Tom, Georgia Goodman, Mohammad Hasdiana, Tony Carnes, Muthiah Vaduganathan, Benjamin Scirica, Jeffrey Schnipper ................................................................. 16

Feasibility of Using Text Messaging to Identify and Assist Patients With Hypertension With Health-Related Social Needs: Cross-Sectional Study (e54530)
Aryn Kormanis, Selina Quinones, Corey Obermiller, Nancy Denizard-Thompson, Deepak Palakshappa ................................................................. 29

Comparing the Efficacy of Targeted and Blast Portal Messaging in Message Opening Rate and Anticoagulation Initiation in Patients With Atrial Fibrillation in the Preventing Preventable Strokes Study II: Prospective Cohort Study (e49590)
Alok Kapoor, Parth Patel, Soumya Chennupati, Daniel Mbusa, Hammad Sadiq, Sanjeev Rampam, Robert Leung, Megan Miller, Kevin Vargas, Patrick Fry, Mary Lowe, Christina Catalano, Charles Harrison, John Catanazzaro, Sybil Crawford, Anne Smith ................................................................. 37

Physical Activity, Heart Rate Variability, and Ventricular Arrhythmia During the COVID-19 Lockdown: Retrospective Cohort Study (e51399)
Sikander Texiwala, Russell de Souza, Suzette Turner, Sheldon Singh ................................................................. 49

User Engagement, Acceptability, and Clinical Markers in a Digital Health Program for Nonalcoholic Fatty Liver Disease: Prospective, Single-Arm Feasibility Study (e52576)
Sigridur Björnsdottir, Hildignurnn Ulfsdottir, Elias Guðmundsson, Kállbrun Sveinsdottir, Ari Ísberg, Bartosz Dobies, Gudlaug Akerlie Magnusdottir, Thrudur Gunnarsdottir, Tekla Karlsdottir, Gudlaug Björnsdottir, Sigurður Sigurðsson, Sæmundur Oddsson, Vilmundur Guðnason ................................................................. 53

Factors That Influence Patient Satisfaction With the Service Quality of Home-Based Teleconsultation During the COVID-19 Pandemic: Cross-Sectional Survey Study (e51439)
Guangxia Meng, Carrie McAiney, Ian McKillop, Christopher Perlman, Shu-Feng Tsao, Helen Chen ................................................................. 68
Cloud-Based Machine Learning Platform to Predict Clinical Outcomes at Home for Patients With Cardiovascular Conditions Discharged From Hospital: Clinical Trial

Phillip C Yang1, MD; Alokkumar Jha1, PhD; William Xu2, BSc; Zitao Song3, MSc; Patrick Jamp4, BSc; Jeffrey J Teuteberg5, MD

1Stanford University School of Medicine, Palo Alto, CA, United States
2Emory University, Atlanta, GA, United States
3North Carolina State University, Raleigh, NC, United States
4Electrical Engineering, University of California, Los Angeles, Mountain View, CA, United States

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

Abstract

Background: Hospitalizations account for almost one-third of the US $4.1 trillion health care cost in the United States. A substantial portion of these hospitalizations are attributed to readmissions, which led to the establishment of the Hospital Readmissions Reduction Program (HRRP) in 2012. The HRRP reduces payments to hospitals with excess readmissions. In 2018, >US $700 million was withheld; this is expected to exceed US $1 billion by 2022. More importantly, there is nothing more physically and emotionally taxing for readmitted patients and demoralizing for hospital physicians, nurses, and administrators.

Given this high uncertainty of proper home recovery, intelligent monitoring is needed to predict the outcome of discharged patients to reduce readmissions. Physical activity (PA) is one of the major determinants for overall clinical outcomes in diabetes, hypertension, hyperlipidemia, heart failure, cancer, and mental health issues. These are the exact comorbidities that increase readmission rates, underlining the importance of PA in assessing the recovery of patients by quantitative measurement beyond the questionnaire and survey methods.

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

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

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

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

(JMIR Cardio 2024;8:e45130) doi:10.2196/45130
KEYWORDS
smart sensor; wearable technology; moving average; physical activity; artificial intelligence; AI

Introduction

Background

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

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

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

Development of an Advanced Data-Driven Analytic Platform

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

After considering various functionalities and requirements, the WD offered the most practical and compliant design solution to monitor discharged patients intelligently. However, the wearable technology for discharged patients should embody different applications and designs specific to their needs. In this paper, we will demonstrate these specifications in more detail. We will present AiCare’s comprehensive technology solution consisting of a WD, Bluetooth low energy (BLE)–enabled iOS infrastructure, an ML algorithm to implement AI in patient care,
and API-enabled web technology to measure the daily activities of patients. In this study, remote data collection, robust XGBoost AI analysis, and the reliable prediction of clinical outcomes are reported.

**Methods**

**Patient Recruitment**

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

**Ethical Considerations**

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

**ML Predictive Platform**

Our comprehensive ML profile of the discharged patients was designed to predict their proper recovery to enable the precision guidance of timely intervention during the 4-week follow-up period. We developed an advanced data-driven XGBoost analytics platform to predict clinical outcomes based on accurate measurements of PA [10]. We compared several techniques to analyze our training data set, including logistic regression, naïve Bayes, support vector machine, and XGBoost. We measured precision, recall, \( F_1 \)-score, area under the ROC curve (AUC-ROC), and average critical activity level, using different data sets. Throughout the analyses, XGBoost provided the highest area under the curve (AUC) values and other measurements.

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

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

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

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

\[
1
\]

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

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

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

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

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

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

\[
P_k = \frac{1}{n} \sum_{i=1}^{n} P_i
\]

(2)
where \(P_k\) represented the data point at time \(k\), and \(n\) was the chosen number of data points. A longer-term SMA was less sensitive in reflecting the change in data movement compared with a shorter-term SMA, which was used to highlight the major trends in time series data. A shorter-term MA was relatively faster to react to changes in trend, which was beneficial to applications that required a QR code. The adjustment of the value \(n\) in the equation measured the different effects of trend analysis. The specific features and the definitions of PA are outlined in Textbox 1.
Textbox 1. Studied features and definitions for physical activity patterns.

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

Kinetic EL

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

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

(3)

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

$$ SVM = \sqrt{a_x^2 + a_y^2 + a_z^2} $$

(4)

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

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

(5)

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

$$ Total \ energy \ level = SVM2(t1) + SVM2(t2) + ... + SVM2(tn) $$

(6)

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

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

$$ Daily \ energy \ level \ difference = EL_{Daytime} - EL_{Nighttime} $$

(7)

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

$$ Active \ percentage(\%) = 100\% - resting \ percentage(\%) $$

(8)

$$ Daily \ active \ percentage \ difference (\%) = Active \ percentage_{Daytime} - Active \ percentage_{Nighttime} $$

(10)

ML Algorithm for PA Analysis

Three key features—daily ELD, normalized ELD, and daily active percentage difference—were used to create the algorithm to predict the possible clinical worsening of discharged patients who demonstrate specific PA patterns. A risk alert was generated when the values of the features that were lower than a specific threshold were detected. The collected data correlated with the detailed measurement of the patients’ PA. We assessed the patients’ quality of recovery through accurate measurements of their activities while they were awake, while performing various activities of daily living (ADL) or participating in PA, and while...
resting. Multiple layers of big data analytics, data mining algorithms, and ML methods were tested. Specifically, we applied the XGBoost ML algorithm. Our solution refined the predictive capability by using the individual PA differences during the active phase (walking, standing, or sitting) versus the resting phase (lying down). XGBoost enhanced the predictive accuracy of healthy recovery versus deterioration at home and determined the need to contact health care professionals. XGBoost distinguished itself from other gradient boost learning methods by using clever penalization of trees, proportional shrinking of leaf nodes, Newton boosting, extra randomization parameter, and the implementation of single distributed systems. These features enabled efficient ML classification of the real-time monitoring of PA to refine the patients’ risk assessment. XGBoost distributed the feed-forward module of PA. The integration started with the PA module of 3 physical acceleration features (1 g, 2 g, and 3 g). The penalty-based system determined the initial risk profiling by PA weight and retrained in a stage-agnostic way to determine the features and penalties to enhance the weight from PA. The final stage repeats the same cycle as stage 2 and provides each patient’s final weight and risk score. As the individual stage is algorithm agnostic, this method provides randomized nonbiased Newtonian analysis. The training data set was achieved by our solution by predicting the clinical outcome based on the individual PA differences during the active phase (walking, standing, or sitting) versus the resting phase (lying down). On the basis of >17,000 person-day data points of 36 participants (unpublished data), we predicted healthy recovery versus death in skilled nursing facility residents based on the PA data and analysis (Figure 2). Using this training data set, our XGBoost algorithm was designed to detect deterioration in the health condition of the discharged patients to generate a risk alert, which suggested the need for early medical intervention by contacting health care professionals, and prevent hospital readmission.

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

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

ROC analysis was used to evaluate a classifier’s prediction performance in biological and medical applications. Each data point in the ROC curve comprises a pair: true-positive rate (sensitivity) and false-positive rate (1−specificity), generated by a discrete classifier with a specific threshold [14]. This study used several periods of MA from 3 days to produce all ROC points in the ROC space. Considering the effect of the imbalanced data set, meaning that the number of healthy discharged patients is greater than the number of discharged
patients classified as deteriorating, we also used the recall-precision curve to evaluate the performance of the proposed algorithms. In the recall-precision curve space, the x-axis and y-axis represent the recall values and precision values, respectively, calculated from the different thresholds.

**User Interface**

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

**Results**

**Patient Enrollment**

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

**Prediction**

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

After the patient’s discharge to home, the AiCare platform is deployed to the patient via the Rockband and iPhone. PA data and trend are displayed on the iPhone app. If there is any negative change in PA, the patient is contacted for a clinical evaluation. The ADL questionnaire is administered to the patient to see whether there is any correlation.
Our solution provided a robust low-cost technology to measure PA and predict clinical outcomes. Our platform also included the nudge technology for an AI-enabled questionnaire. This platform was designed to deliver a seamless, low-cost, and user-friendly environment for the remote monitoring of discharged patients at home to empower the patients and family members. This study analyzed the predictive capability of our platform as described in Textboxes 2-4.

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

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

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

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

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

Questionnaires were also administered at the time of risk alert as well as at the completion of the study (Textbox 7). The ADL questionnaire was administered to augment the PA data. The findings demonstrated modest correlation with the predictive capability. The patients whose condition deteriorated (true-positive group) showed the lowest function in terms of ADL, whereas those who remained stable showed higher response scores (true-negative group). However, there were low ADL scores in the false-positive group and high ADL scores in the false-negative group. The responses to the satisfaction questionnaire demonstrated that this platform was well received. The majority of the users stated that they would recommend the technology to others.
Textbox 3. Formulas of area under the receiver operating characteristic curve metrics (TP=true positive, TN=true negative, FP=false positive, and FN=false negative).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>Accuracy = (\frac{TP + TN}{TP + TN + FP + FN})</td>
</tr>
<tr>
<td><strong>Precision and positive predictive values</strong></td>
<td>Precision = (\frac{TP}{TP + FP}) and Positive predictive value = (\frac{TP}{TP + FN})</td>
</tr>
<tr>
<td><strong>Sensitivity, recall, or true-positive rate (TPR)</strong></td>
<td>Sensitivity = (\frac{TP}{TP + FN})</td>
</tr>
<tr>
<td><strong>Specificity or true-negative rate</strong></td>
<td>Specificity = (\frac{TN}{TN + FP})</td>
</tr>
<tr>
<td><strong>Negative predictive values (NPV)</strong></td>
<td>NPV = (\frac{TN}{TN + FN})</td>
</tr>
<tr>
<td><strong>False-positive rate (FPR)</strong></td>
<td>FPR = 1 – specificity = (\frac{FP}{FP + TN})</td>
</tr>
</tbody>
</table>

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

- Sensitivity: 90.01%
- Specificity: 81.55%
- Positive predictive value: 77.1%
- Negative predictive value: 92.3%
- Accuracy: 85%
- False-positive rate: 18.45%
Textbox 5. Clinical diagnosis and prediction data.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True positive</strong> (n=17)</td>
<td></td>
</tr>
<tr>
<td>Heart failure</td>
<td>9</td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>6</td>
</tr>
<tr>
<td>Atrial</td>
<td>5</td>
</tr>
<tr>
<td>Ventricular</td>
<td>1</td>
</tr>
<tr>
<td>Device</td>
<td>1</td>
</tr>
<tr>
<td>Ischemia</td>
<td>1</td>
</tr>
<tr>
<td><strong>True negative</strong> (n=9)</td>
<td></td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>6</td>
</tr>
<tr>
<td>Atrial</td>
<td>4</td>
</tr>
<tr>
<td>Ventricular</td>
<td>2</td>
</tr>
<tr>
<td>Ischemia</td>
<td>2</td>
</tr>
<tr>
<td>Heart failure</td>
<td>1</td>
</tr>
<tr>
<td><strong>False positive</strong> (n=11)</td>
<td></td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>5</td>
</tr>
<tr>
<td>Atrial</td>
<td>3</td>
</tr>
<tr>
<td>Ventricular</td>
<td>2</td>
</tr>
<tr>
<td>Heart failure</td>
<td>2</td>
</tr>
<tr>
<td>Ischemia</td>
<td>2</td>
</tr>
<tr>
<td>Pulmonary hypertension</td>
<td>2</td>
</tr>
<tr>
<td><strong>False negative</strong> (n=2)</td>
<td></td>
</tr>
<tr>
<td>Arrhythmia, ventricular</td>
<td>1</td>
</tr>
<tr>
<td>Pericarditis</td>
<td>1</td>
</tr>
</tbody>
</table>

Textbox 6. Technical findings.

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

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

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

Discussion

Principal Findings

We screened 249 patients discharged from Stanford Hospital general cardiology and heart failure services. Of these 249 patients, 52 (20.9%) were enrolled, and 36 (14.5%) completed a 28-day analysis looking into the correlation between PA and clinical outcome. Using our XGBoost model, we plotted the true-positive rates against false-positive rates, which helped us to generate the AUC-ROC curve and calculate the 2D AUC value to determine the performance of the model. Our data demonstrated a robust prediction system to forecast the worsening clinical outcomes of these patients based on their PA data, achieving a sensitivity of 87% and a specificity of 79%.
This innovative platform enables low-cost, robust, and precise PA tracking of patients discharged from hospital to predict stable versus unstable clinical recovery, using a WD and an iPhone. This technology is powered by our algorithms, individualized big data, personalized behavior- and function-specific web-based software, and intelligent ML analytics. This comprehensive platform offers an effective convergence of eHealth, AI, and telemedicine technology over internet-enabled mobile devices to leverage the economical, low-cost, and pervasive internet technology and, potentially, may address the socioeconomic divide seen today. Patient care at home by family members or by the individual patient is personalized by AI for maximum safety. This technology will fill an important gap in telemedicine through the use of user-friendly, real-time, and 24/7 remote monitoring for clinical outcome prediction. Patients in transition who are discharged from the hospital, emergency department, or urgent care clinic will benefit from this technology, which can monitor their progress and predict clinical deterioration to enable early intervention for successful recovery at home.

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

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

The evidence-based management of CVDs requires substantial amounts of resources, including advanced therapeutics, complex diagnostics, and sophisticated clinical trials. However, the reliable prediction of clinical outcomes after hospital discharge has presented some challenges [11]. In response, various approaches using ML models such as artificial neural network, decision tree, support vector machine, and naive Bayes have been attempted to predict clinical outcomes, taking into account steps, vital signs, medical conditions, and demographic information [11]. In one of the studies conducted on arrhythmic sudden cardiac death, a deep learning technology approach termed Survival Study of Cardiac Arrhythmia Risk was developed to predict risk for 156 patients with ischemic heart disease. In this model, cardiac magnetic resonance images and covariate data such as demographics, risk factors, electrocardiogram (ECG) measurements, medication use, and outcomes were used as inputs for the 2 branches, where 1 branch is used to visualize the heart’s 3D ventricular geometry, and the other is used to extract arrhythmic sudden cardiac death risk–related imaging features from the cardiac magnetic resonance images. All these data were then used to create a survival curve individualized for each patient with accurate predictions for up to 10 years. However, the limitations of this study include an inability to account for competing risks for the same symptoms and covariates that were not exhaustive or fully inclusive [17]. In another instance, regression and convolutional neural network models were used to predict CVD risk for women. As CVDs are the primary cause of death in women, with evidence of sex bias in the diagnosis of CVDs, the exploration of screening factors for risk detection has never been more urgent. This study assessed the critical risk–screening opportunities offered to women and how the integration of AI can greatly benefit health care providers in interpreting data on women [17,21]. AI may propel the analysis of patient data into meaningful interpretations of patient health, providing health care providers with an additional layer of guidance for patient management plans. After patient discharge, ML is still viable in assisting with remote health monitoring through systems such as Wanda-CVD, which uses patients’ blood pressure and BMI.
measurements as well as their low-density lipoprotein and high-density lipoprotein cholesterol levels to coach them and improve their risk factors for CVD. However, using limited inputs such as blood pressure readings and cholesterol levels may not be entirely meaningful. In a previous study, only less than half of the predictions based solely on cholesterol levels and BMI measurements were correct [8,22].

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

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

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

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

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

Conflicts of Interest
The manuscript presents a potential conflict of interest due to the financial involvement of its authors, WX’s family member, PCY, ZS, AJ, and PJ in AiCare, the company responsible for sponsoring the clinical trials discussed in the manuscript. The conflict arises from the authors holding shares in AiCare, indicating a direct financial interest in the success and promotion of
the company’s products. Full disclosure and transparency regarding these financial relationships are essential for maintaining the integrity of the scientific work and ensuring the reader’s ability to assess potential biases.

References

Abbreviations

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

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

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

©Phillip C Yang, Alokkumar Jha, William Xu, Zitao Song, Patrick Jamp, Jeffrey J Teuteberg. Originally published in JMIR Cardio (https://cardio.jmir.org), 01.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cardio, is properly cited. The complete bibliographic information, a link to the original publication on https://cardio.jmir.org, as well as this copyright and license information must be included.
Formative Perceptions of a Digital Pill System to Measure Adherence to Heart Failure Pharmacotherapy: Mixed Methods Study

Peter R Chai1,2,3,4, MD, MS; Jenson J Kaithamattam1, BS; Michelle Chung1, BA; Jeremiah J Tom1, MBBS; Georgia R Goodman1,4, BS; Mohammad Adrian Hasdienda1, MD, MSc, MMSc; Tony Christopher Carnes5, PhD; Muthiah Vaduganathan6, MD, MPH; Benjamin M Scirica6, MD, MPH; Jeffrey L Schnipper7, MD

1Department of Emergency Medicine, Brigham and Women’s Hospital, Boston, MA, United States
2Department of Psychosocial Oncology and Palliative Care, Dana Farber Cancer Institute, Boston, MA, United States
3The Koch Institute for Integrated Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, United States
4The Fenway Institute, Boston, MA, United States
5eTectRx, Gainesville, FL, United States
6Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women’s Hospital, Boston, MA, United States
7Division of Hospital Medicine, Department of Medicine, Brigham and Women’s Hospital, Boston, MA, United States

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

Abstract

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

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

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

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

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

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

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

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

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

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

**Methods**

**Participants**

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

**Procedures**

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

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

**Measures**

**Qualitative Interview**

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

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

aSGLT2i: sodium-glucose co-transporter 2 inhibitors.
bDPS: digital pill system.

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

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

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

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

Results
Participant Characteristics
Over the study period, 96 individuals met the inclusion criteria. Of these, 43 (45%) were discharged before they could be
approached by the study team. Of the remaining 53 individuals, 12 (23%) were unavailable for consent, and 21 (40%) declined to participate. The reasons provided for declining participation included the time commitment for study procedures (n=2), general lack of interest (n=10), lack of knowledge of current medications (n=1), perception that they did not match the target study population (n=1), dissatisfaction with current clinical care (n=1), and reason unknown (n=6). A total of 20 participants consented and completed all study procedures (mean age 68, SD 14.3 years). The sample was predominantly female (n=11, 55%), White (n=13, 65%), and non-Hispanic (n=18, 90%). Full sociodemographic information is provided in Table 2.

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

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

$GED$: general educational development.

Quantitative Results

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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HF status</strong></td>
<td></td>
</tr>
<tr>
<td>Duration of HF (years), n (%)</td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>1 (5)</td>
</tr>
<tr>
<td>1-2</td>
<td>2 (10)</td>
</tr>
<tr>
<td>2-5</td>
<td>6 (30)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>11 (55)</td>
</tr>
<tr>
<td>Primary physician managing HF treatment, n (%)</td>
<td></td>
</tr>
<tr>
<td>Cardiologist</td>
<td>13 (65)</td>
</tr>
<tr>
<td>Primary care physician</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Does not know</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Number of prescribed HF medications, n (%)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1 (5)</td>
</tr>
<tr>
<td>2-5</td>
<td>10 (50)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>9 (45)</td>
</tr>
<tr>
<td>Type of HF, n (%)</td>
<td></td>
</tr>
<tr>
<td>HFpEF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10 (50)</td>
</tr>
<tr>
<td>HFrEF&lt;sup&gt;b&lt;/sup&gt;</td>
<td>10 (50)</td>
</tr>
<tr>
<td>Number of hospital admissions for HF over the last 12 months, n (%)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2 (10)</td>
</tr>
<tr>
<td>1</td>
<td>7 (35)</td>
</tr>
<tr>
<td>2-5</td>
<td>10 (50)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Number of physician encounters due to concerns surrounding worsening HF over last 12 months, n (%)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>7 (35)</td>
</tr>
<tr>
<td>1-5</td>
<td>10 (50)</td>
</tr>
<tr>
<td>6-10</td>
<td>1 (5)</td>
</tr>
<tr>
<td>11-20</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Degree of concern about HF, n (%)</td>
<td></td>
</tr>
<tr>
<td>Slightly concerned</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Moderately concerned</td>
<td>5 (25)</td>
</tr>
<tr>
<td>Very concerned</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Extremely concerned</td>
<td>9 (45)</td>
</tr>
<tr>
<td>Pharmacotherapy adherence</td>
<td></td>
</tr>
<tr>
<td>Percentage of self-reported HF medication adherence over last 3 months, mean (SD)</td>
<td>90.1 (17.1)</td>
</tr>
<tr>
<td>Uses a system to maintain medication adherence, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>12 (60)</td>
</tr>
<tr>
<td>No</td>
<td>8 (40)</td>
</tr>
<tr>
<td>Medication adherence systems used, n (%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Smart pill box</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Pill organizer</td>
<td>10 (83)</td>
</tr>
</tbody>
</table>
Visualization of adherence patterns would motivate medication adherence, n (%)  
Yes: 151 (55)  
No: 7 (35)  
Unsure: 2 (10)  

Willingness to use the DPS\textsuperscript{d}, n (%)  
Yes: 12 (60)  
No: 8 (40)  

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

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

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

\textsuperscript{a}Participants were provided with the opportunity to select multiple options.

### Qualitative Results

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

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

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

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

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

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

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

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

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

**Perceptions Around Privacy and Sharing of DPS Data**

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

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

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

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

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

**Responses to DPS Messaging Components**

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

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

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

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

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

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

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

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

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

Suggested Improvements for Future Iterations

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

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

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

Discussion

Overview

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

Principal Findings

After learning about the DPS, 60% (12/20) of participants perceived the system as an “acceptable” strategy to measure HF pharmacotherapy adherence. Participants also expressed the “usefulness” of the DPS based on its perceived ability to motivate adherence, provide accountability, and avoid double-dosing. For most individuals, having incontrovertible evidence of their adherence (or nonadherence)—especially in the context of clinical care, where their DPS adherence data could aid discussions with their HF physicians and guide future medication decisions—was perceived as the most valuable benefit. These qualitative findings reinforce other proposed
benefits in the literature of leveraging real-time adherence data to not only address medication adherence in chronic disease but potentially enhance the patient-physician relationship by providing key data to ground conversations surrounding disease progression [34]. This concept was reflected in the quantitative portion of the study, where 55% (11/20) of individuals considered having a visual record of ingestion patterns over time as a motivating factor to continue to maintain medication adherence. These perceptions are consistent with other investigations that suggest individuals with other chronic diseases find value in DPS-based adherence data as a technique to guide pharmacotherapy [35-37]. Together, our data suggests that future research investigations should seek to understand the feasibility of real-world DPS operation among individuals with HF, as well as evaluate the impact of adherence metrics on disease progression and treatment regimens in both research and clinical deployment contexts.

Efforts to optimize the design of DPS technology to measure adherence to HF medication should also include a customizable messaging architecture that responds to detected patterns of adherence. Based on our data, messaging components should include confirmation messages to help individuals recognize that they correctly operated the DPS and recorded their medication ingestion. Most importantly, messaging modules should include nonadherence reminder messages that respond to adherence patterns from the DPS, which participants identified as a critical and valuable component of the system. A major theme that emerged from our interviews was participants’ desire for control over both the timing and content of reminder messages related to their adherence patterns. While some wanted daily or even more frequent messaging that would coach them through adherence lapses, others preferred only on-demand access to their adherence data and less frequent feedback from the system. Importantly, some participants also reported that reminder messages could represent a potential method for providing educational information about HF and reinforcing DPS users’ understanding of the consequences of medication nonadherence. These emerging themes demonstrate the importance of involving patients in the design and delivery of adherence interventions linked to digital health systems such as the DPS [29,38]. Barriers to the use of the DPS included discomfort with technology among some users and concerns about the privacy and security of their data. Ultimately, participants viewed the DPS in its current iteration as usable, but they suggested several key improvements that would enable better integration into daily life. Some of these suggestions, including miniaturizing the Reader and providing alternative off-body systems that can collect adherence data, are currently under investigation in other ingestible sensor trials [39]. Additionally, participants expressed that DPS deployment should only occur alongside a detailed discussion with users about the safety and security of the system. While the DPS is FDA 510k cleared, participants emphasized the importance of providing users with data from past users of the system, particularly surrounding any DPS-related adverse events [25].

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

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

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

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

Authors’ Contributions
PRC, JLS, MV, and BMS designed the study and developed the qualitative interview guide. JJK and MC conducted participant screening and approach, enrolment, and qualitative interviews. JJK, JTT, and GRG generated the codebook for evaluation of
results and facilitated the coding using NVivo. JJK, JJT, PRC, and GRG evaluated and identified major themes in the qualitative results. All authors read and approved the final manuscript.

Conflicts of Interest

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

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.

[PDF File (Adobe PDF File), 417 KB - cardio_v8i1e48971_app1.pdf ]

Multimedia Appendix 2
Qualitative interview guide.

[DOCX File, 27 KB - cardio_v8i1e48971_app2.docx ]

References


34. Richey AG, Kovacs I, Browne S. Use of an ingestible, sensor-based digital adherence system to strengthen the therapeutic relationship in serious mental illness. JMIR Ment Health 2022;9(12):e39047 [FREE Full text] [doi: 10.2196/39047] [Medline: 36459392]


Abbreviations

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

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

Please cite as:

URL: https://cardio.jmir.org/2024/1/e48971
PMID: 38358783

©Peter R Chai, Jenson J Kaithamattam, Michelle Chung, Jeremiah J Tom, Georgia R Goodman, Mohammad Adrian Hasdianda, Tony Christopher Carnes, Muthiah Vaduganathan, Benjamin M Scirica, Jeffrey L Schnipper. Originally published in JMIR Cardio (https://cardio.jmir.org), 15.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cardio, is properly cited. The complete bibliographic information, a link to the original publication on https://cardio.jmir.org, as well as this copyright and license information must be included.
Feasibility of Using Text Messaging to Identify and Assist Patients With Hypertension With Health-Related Social Needs: Cross-Sectional Study

Aryn Kormanis¹, DO; Selina Quinones¹, BS; Corey Obermiller¹, MAS; Nancy Denizard-Thompson¹, MD; Deepak Palakshappa¹,²,³, MS, MD

¹Department of Internal Medicine, Wake Forest University School of Medicine, Winston Salem, NC, United States
²Department of Pediatrics, Wake Forest University School of Medicine, Winston Salem, NC, United States
³Department of Epidemiology and Prevention, Wake Forest University School of Medicine, Winston Salem, NC, United States

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

Abstract

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

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

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

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

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

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

https://cardio.jmir.org/2024/1/e54530
KEYWORDS
social determinants of health; health-related social needs; mobile health; health information technology; feasibility; mobile phone; SMS text messaging; message; pilot study; patients; patient; hypertension; screening

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

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

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

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

SMS Text Messages
We developed the SMS text message–based on a detailed literature review of the social needs literature (Multimedia Appendix 1). The AHWFB digital communication committee also reviewed the SMS text messages. The committee includes patients, clinicians, hospital administrators, and members of the institutional review board. The committee reviews all SMS text messages that may be sent to patients for either research or clinical purposes at the institution, and they provide input on the wording of the messages. We also reviewed the SMS text messages with patients who were not included in the study to assess face validity and to provide additional input on the messages. The SMS text message included a link to a questionnaire with 6 questions from the CMS Accountable Health Communities Health-Related Social Needs Screening Tool [7]. The 6 questions included 2 questions about housing (including 1 about patients’ current housing situation and 1 about problems with housing), 2 questions about food insecurity, 1 question about transportation, and 1 question about use. We limited to questions regarding food, housing, transportation,
and utilities because there were local resources available to assist patients with these domains. The initial message was sent to the patient’s cell phone number listed in the EHR on March 14, 2023. All of the 4446 patients eligible had a phone number listed in the EHR. For those that did not respond, the same message was sent again 1 week later. Responses were documented in the EHR. The SMS text message and questionnaire were sent in English or Spanish based on the preferred language of the patient listed in the EHR. We used the standard scoring to identify patients who screened positive for a social need. The patients were identified as having housing instability if they provided a response other than “I have a steady place to live” to the first housing question or if they provided a response other than “None of the above” to the second question. Patients were identified as having food insecurity if they responded to either of the 2 questions with “sometimes true” or “often true.” Patients were identified as having a lack of transportation if they responded “yes” to the transportation question, and they were identified as having difficulty with utilities if they responded “yes” or “already shut off” to the utilities question.

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

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

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

Of the 200 patients, the SMS text message was unable to be delivered to 9 patients (either because the number listed was no longer working or was not a cell phone). A total of 17 (8.5%) patients completed the social needs questionnaire (Table 1). We did not find a significant difference in demographics between patients who did and those who did not complete the questionnaire. Of the 17 who completed the questionnaire, 5 (29.4%) reported at least 1 unmet social need, but only 2 chose to receive local resource information. One person completed the follow-up questionnaire and reported that they learned about new community resources through the process.
### Table 1. Study patient characteristics\(^a\).

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

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

\(^b\)N/A: not applicable.
Discussion

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

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

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

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

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

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

Acknowledgments

DP is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under grant (K23HL146902). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
of Health. The funding organization has no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the study; and decision to submit the study for publication. DP reports personal fees from WellCare of North Carolina outside of the submitted work.

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

Conflicts of Interest
None declared.

Multimedia Appendix 1
Data and questions included in text messages.

References


Abbreviations

AHWF B: Atrium Health Wake Forest Baptist Health
CMS: Centers for Medicare and Medicaid
EHR: electronic health record
RE-AIM: Reach, Effectiveness, Acceptability, Implementation, and Maintenance
©Aryn Kormanis, Selina Quinones, Corey Obermiller, Nancy Denizard-Thompson, Deepak Palakshappa. Originally published in JMIR Cardio (https://cardio.jmir.org), 13.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cardio, is properly cited. The complete bibliographic information, a link to the original publication on https://cardio.jmir.org, as well as this copyright and license information must be included.
Comparing the Efficacy of Targeted and Blast Portal Messaging in Message Opening Rate and Anticoagulation Initiation in Patients With Atrial Fibrillation in the Preventing Preventable Strokes Study II: Prospective Cohort Study

Alok Kapoor1,2, MS, MD; Parth Patel1, BS; Soumya Chennupati1, BS; Daniel Mbusa1, BS; Hammad Sadiq1, BS; Sanjeev Rampam1, BS; Robert Leung1,2, MBA, MD; Megan Miller3, BS; Kevin Rivera Vargas3, BS; Patrick Fry4, BS; Mary Martin Lowe5, PhD; Christina Catalano4, MBA; Charles Harrison4, MD; John Nicholas Catanzaro4, MD; Sybil Crawford1, PhD; Anne Marie Smith6, MBA

1University of Massachusetts Chan Medical School, Worcester, MA, United States
2University of Massachusetts Memorial Health Care, Worcester, MA, United States
3College of Pharmacy, University of Florida, Jacksonville, FL, United States
4College of Medicine, University of Florida, Jacksonville, FL, United States
5Learning Advisors LLC, Chicago, IL, United States
6Heart Rhythm Society, Washington, DC, United States

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

Abstract

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

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

Methods: Using the 2 approaches, we sent patient portal messages to patients with AF and grouped patients by high-risk patients on anticoagulation (group 1), high-risk patients off anticoagulation (group 2), and low-risk patients who may become eligible for anticoagulation in the future (group 3). Risk was classified based on the congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age between 65 and 74 years, and sex category (CHA2DS2-VASc) score. The messages contained a link to the Upbeat website of the Heart Rhythm Society, which displays print and video materials about AF and anticoagulation. We then tracked message opening, review of the website, anticoagulation use, and administered patient surveys across messaging approaches and sites using Epic Systems (Epic Systems Corporation) electronic health record data and Google website traffic analytics. We then conducted chi-square tests to compare potential differences in the proportion of patients opening messages and other evaluation metrics, adjusting for potential confounders. All statistical analyses were performed in SAS (version 9.4; SAS Institute).

Results: We sent 1686 targeted messages and 1450 blast messages. Message opening was significantly higher with the targeted approach for patients on anticoagulation (723/1156, 62.5% vs 382/668, 57.2%; P=.005) and trended the same in patients off
anticoagulation; subsequent website reviews did not differ by messaging approach. More patients off anticoagulation at baseline started anticoagulation with the targeted approach than the blast approach (adjusted percentage 9.3% vs 2.1%; \( P < .001 \)).

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

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

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

Introduction

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

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

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

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

Methods

Overview

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

Study Design

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

Setting

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

Participants

We included patients aged 18 years or older with AF with active MyChart patient portal accounts and who had at least 1 office visit in the 12 months before the start of our messaging intervention in November 2021. At UMass, starting each workday from November 2021 to February 2022, we ran the Epic System’s Reporting Workbench that identified patients based on their having an appointment (office or tele-visit type) with a primary or cardiology care provider scheduled to take place within a week. At UFL, from November 2021 to
December 2021, we identified patients based on a previously established registry of those patients with AF who had a visit with a primary or cardiology care provider in the previous year. At both sites, we grouped patients based on their anticoagulation status (eg, on or off anticoagulation) and congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, stroke, vascular disease, age 65-74 years, and sex scale (CHA2 DS2-VASc) score. Specifically, group 1 included those at high risk (eg, CHA2 DS2-VASc score of ≥2 for men and ≥3 for women) and currently on anticoagulation; group 2 included those at high risk and off anticoagulation; and group 3 included those at low risk (eg, CHA2 DS2-VASc score <2 for men and <3 for women) and not on anticoagulation.

**Outcomes, Variables, or Data Sources**

**Message Opening**

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

**Website Review**

Using Google Analytics (Google LLC), we tracked the number of unique page views as the value representing the total number of unique sessions. As patients may have received more than 1 message throughout the study (corresponding to 2 separate visits or in the case of canceled and rescheduled visits), we selected the number of messages sent as the denominator. We then calculated the percentage of messages resulting in a unique page view, with the number of unique page views as the numerator and compared this across sites. We also compared the “bounce rate” across sites, which represents the percentage of all sessions on a site in which users only viewed a single page. Google documentation [18] notes that bounce rates should be interpreted within the context of a specific website’s purpose. Upbeat, the website our messages directed patients toward, has many links within the context of a specific website’s purpose. Upbeat, the documentation [18] notes that bounce rates should be interpreted within the context of a specific website’s purpose.

**Independent Exposure**

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

**Other Exposures: Anticoagulation Outcome Only**

We included stroke risk based on the CHA2 DS2-VASc score, which is comprised of congestive heart failure, hypertension, age, diabetes, previous stroke, vascular disease, and gender. To adjust further for potential confounders of the association between anticoagulation use and message opening, we included demographics omitted in that score (ie, race, ethnicity, language preference, and primary insurance). Finally, we included chronic kidney disease and anemia. In general, we relied on the International Classification of Diseases, tenth edition codes for the presence of a comorbid condition. For chronic kidney disease, low platelet count, and anemia, we relied on laboratory data.

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

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

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

**Anticoagulation Outcome**

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

To address potential bias from the confounder of the difference in populations at the 2 different sites, we computed the adjusted percentage of patients on anticoagulation between messaging approaches. More specifically, we constructed a generalized logistic mixed model with anticoagulation status (on or off) as the dependent variable and messaging approach (targeted vs blast) as the independent variable. We also included a random effect for provider to account for potential clustering and several covariates to adjust for potential confounders of anticoagulation. Covariates included variables making it more likely to be on anticoagulation (eg, higher stroke risk expressed through the CHA$_2$DS$_2$-VASc score) as well as factors making it less likely to be on anticoagulation (eg, anemia, chronic kidney disease, and high BMI). We did this separately for groups 1 and 2.

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

**Ethical Considerations**

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

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

**Results**

**Message Opening**

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

Message opening was moderately high at both sites and across groups, with the highest opening rates in group 1 (723/1156, 62.5%) at UMass and group 3 (87/150, 57.3%) at UFL. Message opening in group 1 was significantly higher at UMass than at UFL (723/1156, 62.5% vs 382/668, 57.2%; $P=.005$). We did not find a statistically significant difference in message opening rates between the targeted (UMass) and blast (UFL) messaging approaches for group 2 (274/438, 62.6% vs 335/632, 53%; $P=.09$) and group 3 (52/92, 56.5% vs 86/150, 57.3%; $P=.22$).
### Table 1. Key characteristics in patients receiving messages with targeted versus blast approach (percentages may not sum to 100 due to rounding).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Targeted messaging (University of Massachusetts)</th>
<th>Blast messaging (University of Florida)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
<td><strong>1 (high risk on anticoagulation; n=1156), n (%)</strong></td>
<td><strong>2 (high risk off anticoagulation; n=438), n (%)</strong></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65</td>
<td>272 (23.5)</td>
<td>96 (21.9)</td>
</tr>
<tr>
<td>65-74</td>
<td>374 (32.4)</td>
<td>140 (32)</td>
</tr>
<tr>
<td>≥75</td>
<td>501 (43.3)</td>
<td>196 (44.8)</td>
</tr>
<tr>
<td>Missing</td>
<td>9 (0.8)</td>
<td>6 (1.4)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>436 (37.7)</td>
<td>186 (42.5)</td>
</tr>
<tr>
<td>Male</td>
<td>708 (61.2)</td>
<td>246 (56.2)</td>
</tr>
<tr>
<td>Missing</td>
<td>12 (1)</td>
<td>6 (1.4)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>18 (1.6)</td>
<td>4 (0.9)</td>
</tr>
<tr>
<td>Other</td>
<td>49 (4.2)</td>
<td>19 (4.3)</td>
</tr>
<tr>
<td>White</td>
<td>1080 (93.4)</td>
<td>414 (94.5)</td>
</tr>
<tr>
<td>Decline to answer, missing, or unknown</td>
<td>9 (0.8)</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>44 (3.8)</td>
<td>14 (3.2)</td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>1089 (94.2)</td>
<td>420 (95.9)</td>
</tr>
<tr>
<td>Decline to Answer</td>
<td>22 (1.9)</td>
<td>4 (0.9)</td>
</tr>
<tr>
<td>Unknown or missing</td>
<td>1 (0.1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Language preference</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>1109 (95.9)</td>
<td>419 (95.7)</td>
</tr>
<tr>
<td>Not English</td>
<td>47 (4.1)</td>
<td>19 (4.3)</td>
</tr>
<tr>
<td>Unknown or missing</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

### Website Review

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

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

### Survey-Based Outcomes

From Group 1, 93 and 59 patients answered our survey using the targeted and blast messaging approaches, respectively. There was not a significant difference in patient reports of discussion with their provider about stroke risk, the duration of current anticoagulation use, or the frequency of missing doses of anticoagulation. Notably, forgetfulness and other reasons (apart from costs, side effects, or lack of benefit) comprise the majority of reasons for forgetting doses across messaging approaches. Most patients in both the targeted vs blast messaging groups strongly agreed or agreed that the materials from the HRS were easy-to-understand (68/82, 83% vs 13/15, 87%), were useful (69/82, 84% vs 14/15, 93%), as well as something they would recommend (71/83, 85% vs 14/15, 93%) without any of the differences reaching statistical significance (Table S2 in Multimedia Appendix 4 [27] for details).
From Group 2, a total of 9/25 patients answered our survey using the targeted and blast messaging approaches, respectively. More patients in the UMass group had discussed their stroke risk with their physician at UMass (16/25, 64% vs 3/9, 33%; P=.04). Among the patients in the targeted approach, only 26% (6/25) reported concern about the risk of bleeding as a cause for stopping anticoagulation. Only 4 patients from the blast approach answered this item, limiting comparison. The majority of patients strongly agreed or agreed that the materials from the HRS were easy-to-understand and useful, as well as something they would recommend (also, comparisons were limited due to only 3 patients from the blast messaging group answering this item; Table S3 in Multimedia Appendix 4 [27]).

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

**Anticoagulation**

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

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

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

^aBlack includes Black of African American and multiracial, including Black or African American, Hispanic includes those individuals reporting Hispanic or Latino ethnicity, Asian includes White Hispanic. There were no individuals who reported Black race or Hispanic ethnicity. Other include Asian, Native American, Alaska Native, and others.
was more effective than blast messaging in achieving message opening for those on anticoagulation. There was a trend toward increased message opening among patients off anticoagulation (274/438, 62.6% for the targeted approach versus 335/632, 53% for the blast approach). This increased message opening may have explained some anticoagulation starts, but replication at other sites would be helpful in drawing firm conclusions. Given the low rates of website reviews it is unlikely that it contributed to anticoagulation starts and would not be valuable to include in future programs, at least in the way we delivered it (as a simple website link). Education provided directly in the message or within the health portal is likely to be more effective than requesting patients to review external websites.

There are other implications for our findings. The best approach to messaging patients should also factor in local resources. Our approach to sending targeted messaging required the daily execution of a workbench report and subsequent filtering and transmission of portal messages. In the future, we anticipate that we could automate the manual steps and link the messaging to portal messages sent to patients related to preparation for anticipated health programs such as yearly vaccination campaigns. As a result, the best approach for institutions and providers to take when determining how to deliver messages to their patients.

Limitations

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

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

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted percentage on anticoagulation</th>
<th>Adjusted percentage on anticoagulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Targeted messaging (UMass(^a)), n (%)</td>
<td>Blast messaging (UFL(^b)), n, (%)</td>
</tr>
<tr>
<td>Group 1 (high risk, on anticoagulation at baseline)</td>
<td>438 (95.4)</td>
<td>179 (96.2)</td>
</tr>
<tr>
<td>Group 2 (high risk, off anticoagulation at baseline)</td>
<td>21 (11.9)</td>
<td>3 (3.0)</td>
</tr>
</tbody>
</table>

\(^a\)UMass: University of Massachusetts.

\(^b\)UFL: University of Florida.

\(^c\)Chi-square–based P value.

\(^d\)Only percentages are shown.

\(^e\)Value derived from generalized estimating equation adjusting for age, gender, BMI, patient race-ethnicity, insurance, CHA\(_2\)DS\(_2\)-VASc score, presence of anemia (ie, hemoglobin <13 g/dL for male candidates, <12 g/dL for female candidates, and level of chronic kidney disease).

Discussion

Principal Results

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

Comparison With Previous Work

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

The main implication of this study is that targeted messaging was more effective than blast messaging in achieving message opening closer to the time of the visit would achieve the optimal timing within 6 months for sending a message. Presumably closer to the time of the visit would achieve the best results.

The authors do not further report the optimal timing within 6 months for sending a message. Presumably closer to the time of the visit would achieve the optimal timing within 6 months for sending a message.
opening and website review, where we did not have patient-level variables. For the anticoagulation outcome, we adjusted for known confounders of the use of anticoagulation, including demographics, stroke risk score, and bleeding risk factors (ie, anemia and chronic kidney disease). Many other factors, including other indications for anticoagulation, type of anticoagulant, baseline health literacy, and computer literacy, may, however, have been different across our sites. In addition, because the site of care dictated the receipt of one versus the other messaging approach and we had limited information about the reason for receiving care at one versus the other site, we did not pursue propensity or other causal inference modeling approaches. Other institutional-based programs may have explained the increase with the targeted approach. A human-centered design approach has successfully overcome limitations in other messaging programs cited in the literature [32,36]. Oake et al [36] observed that an automated voice messaging response system for communicating anticoagulation testing and dosage schedules to patients led to improved anticoagulation monitoring. Another limitation was that we were not able to ascertain if patients read our message, only that our portal message was opened. In many cases, a family member will be opening the message. Website review and survey responses may be limited in the same way. Although education by proxy through a family member may lead to decisions to take anticoagulation or stay on it, we were not able to distinguish the discrete effect of direct versus proxy communication in the current study. Our results may also not generalize to non-White populations, which is significant given the lower adherence of non-Whites [8]. Lastly, it is important to note the impact of COVID-19 and the timing of the UFL messages on the project. UFL messages were sent out in December, with a follow-up in January. In addition to patients receiving holiday-related emails, COVID-19 was surging as well. It is unclear how these two variables may have impacted message opening.

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


18. Bounce rate. Google Analytics. URL: https://support.google.com/analytics/answer/1009409?hl=en#xt=t=textsAbout%20bounce%20rate&text=Bounce%20rate%20is%20single%2Fpage%20request%20to%20%20Analytics%20server [accessed 2023-12-28]

https://cardio.jmir.org/2024/1/e49590 JMIR Cardio 2024 | vol. 8 | e49590 | p.46 (page number not for citation purposes)
19. What is a good bounce rate? And what's a bad one? Fullstory Education Team. 2023. URL: https://www.fullstory.com/blog/what-is-a-good-bounce-rate/ [accessed 2023-12-28]


26. The difference between Google ads clicks, and sessions, users, entrances, pageviews, and unique pageviews in analytics. Google Analytics. URL: https://support.google.com/analytics/answer/1257084?hl=en#zipp [accessed 2023-10-23]


---

**Abbreviations**

AF: atrial fibrillation  
CHA2DS2-VASc: congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age 65-74 years, sex scale  
EHR: electronic health record  
HRS: Heart Rhythm Society  
IRB: institutional review board  
UFL: University of Florida College of Medicine-Jacksonville  
UMass: University of Massachusetts Chan Medical School
Comparing the Efficacy of Targeted and Blast Portal Messaging in Message Opening Rate and Anticoagulation Initiation in Patients With Atrial Fibrillation in the Preventing Preventable Strokes Study II: Prospective Cohort Study


© Alok Kapoor, Parth Patel, Soumya Chennupati, Daniel Mbusa, Hammad Sadiq, Sanjeev Rampam, Robert Leung, Megan Miller, Kevin Rivera Vargas, Patrick Fry, Mary Martin Lowe, Christina Catalano, Charles Harrison, John Nicholas Catanzaro, Sybil Crawford, Anne Marie Smith. Originally published in JMIR Cardio (https://cardio.jmir.org), 24.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cardio, is properly cited. The complete bibliographic information, a link to the original publication on https://cardio.jmir.org, as well as this copyright and license information must be included.
Physical Activity, Heart Rate Variability, and Ventricular Arrhythmia During the COVID-19 Lockdown: Retrospective Cohort Study

Sikander Z Texiwala¹, MD; Russell J de Souza²,³, RD, ScD; Suzette Turner¹,⁴, BScN; Sheldon M Singh¹,⁵, MD

¹Schulich Heart Center, Sunnybrook Health Sciences, Toronto, ON, Canada
²Department of Health Research Methods, Evidence, and Impact, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
³Population Health Research Institute, Hamilton Health Sciences Corporation, Hamilton, ON, Canada
⁴Lawrence Bloomberg Faculty of Nursing, University of Toronto, ON, Canada
⁵Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

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

Abstract

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

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

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

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

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

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

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

Introduction

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

Patients with implantable cardioverter defibrillators (ICDs) are at risk for VA. ICDs contain sensors that can quantify the physical activity of a patient who has an ICD implanted. Acute
increases in physical activity can increase the risk of VA [6]. ICDs monitor the changes in the patient’s heart rate. Heart rate variability (HRV) summarizes the beat-to-beat changes in heart rate and reflects the balance between the sympathetic and parasympathetic nervous system. A reduction in HRV can occur during times of stress due to increased sympathetic activation. A reduction in HRV predicts VA [7]. Assessing the changes in physical activity and HRV may provide insight into the lack of increased VA observed during the onset of the COVID-19 lockdown.

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

**Methods**

**Study Cohort**

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

**Study Periods**

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

**VA, Physical Activity, and Heart Rate Variability**

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

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

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

**Statistical Analysis**

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

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

**Ethical Considerations**

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

**Results**

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

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

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

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

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

**Discussion**

**Principal Findings**

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

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

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

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

Conclusions

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

Acknowledgments

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

Data Availability

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

Authors’ Contributions

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

Conflicts of Interest

None declared.

References


8. Huang JL, Chio CT, Chen YT, Chen SA. Sudden changes in heart rate variability during the 1999 Taiwan earthquake. Am J Cardiol 2001;87(2):245-248, A9. [doi: (10.1016/s0002-9149(00)01331-x) [Medline: 11152854]

Abbreviations

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

©Sikander Z Texiwala, Russell J de Souza, Suzette Turner, Sheldon M Singh. Originally published in JMIR Cardio (https://cardio.jmir.org), 05.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cardio, is properly cited. The complete bibliographic information, a link to the original publication on https://cardio.jmir.org, as well as this copyright and license information must be included.
Original Paper

User Engagement, Acceptability, and Clinical Markers in a Digital Health Program for Nonalcoholic Fatty Liver Disease: Prospective, Single-Arm Feasibility Study

Sigridur Björnsdottir¹, MD, PhD; Hildigunnur Ulfsdottir², MPH, MD; Elias Freyr Gudmundsson², BSc, MSc; Kolbrún Sveinsdottir², MSc; Ari Pall Isberg², MSc; Bartosz Dobies³, MSc; Guðlaug Erla Akerlie Magnusdottir², BSc; Thrúður Gunnarsdóttir², PhD; Tekla Karlsdóttir³, MD; Guðlaug Björnsdóttir³,⁴, MSc; Sigurdur Sigurðsson³,⁴, PhD; Saemundur Oddsson², MD; Vilmundur Gudnason³,⁴, MD, PhD

¹Department of Endocrinology, Metabolism and Diabetes, Karolinska Institutet, Stockholm, Sweden
²Sidekick Health, Kopavogur, Iceland
³Icelandic Heart Association, Kopavogur, Iceland
⁴School of Health Sciences, Faculty of Medicine, University of Iceland, Reykjavik, Iceland

Corresponding Author:
Sigridur Björnsdottir, MD, PhD
Department of Endocrinology, Metabolism and Diabetes
Karolinska Institutet
Solnavägen 1
Stockholm, 171 77
Sweden
Phone: 46 8 524 800 00
Email: sigridur.bjornsdottir@ki.se

Abstract

Background: Nonalcoholic fatty liver disease (NAFLD) has become the most common chronic liver disease in the world. Common comorbidities are central obesity, type 2 diabetes mellitus, dyslipidemia, and metabolic syndrome. Cardiovascular disease is the most common cause of death among people with NAFLD, and lifestyle changes can improve health outcomes.

Objective: This study aims to explore the acceptability of a digital health program in terms of engagement, retention, and user satisfaction in addition to exploring changes in clinical outcomes, such as weight, cardiometabolic risk factors, and health-related quality of life.

Methods: We conducted a prospective, open-label, single-arm, 12-week study including 38 individuals with either a BMI >30, metabolic syndrome, or type 2 diabetes mellitus and NAFLD screened by FibroScan. An NAFLD-specific digital health program focused on disease education, lowering carbohydrates in the diet, food logging, increasing activity level, reducing stress, and healthy lifestyle coaching was offered to participants. The coach provided weekly feedback on food logs and other in-app activities and opportunities for participants to ask questions. The coaching was active throughout the 12-week intervention period. The primary outcome was feasibility and acceptability of the 12-week program, assessed through patient engagement, retention, and satisfaction with the program. Secondary outcomes included changes in weight, liver fat, body composition, and other cardiometabolic clinical parameters at baseline and 12 weeks.

Results: In total, 38 individuals were included in the study (median age 59.5, IQR 46.3-68.8 years; n=23, 61% female). Overall, 34 (89%) participants completed the program and 29 (76%) were active during the 12-week program period. The median satisfaction score was 6.3 (IQR 5.8-6.7) of 7. Mean weight loss was 3.5 (SD 3.7) kg (P<.001) or 3.2% (SD 3.4%), with a 2.2 (SD 2.7) kg reduction in fat mass (P<.001). Relative liver fat reduction was 19.4% (SD 23.9%). Systolic blood pressure was reduced by 6.0 (SD 13.5) mmHg (P=.009). The median reduction was 0.14 (IQR 0-0.47) mmol/L for triglyceride levels (P=.003), 3.2 (IQR 0-5.4) µU/ml for serum insulin (s-insulin) levels (P=.003), and 0.5 (IQR –0.7 to 3.8) mmol/mol for hemoglobin A₁c (HbA₁c) levels (P=.03). Participants who were highly engaged (ie, who used the app at least 5 days per week) had greater weight loss and liver fat reduction.

Conclusions: The 12-week-long digital health program was feasible for individuals with NAFLD, receiving high user engagement, retention, and satisfaction. Improved liver-specific and cardiometabolic health was observed, and more engaged participants
showed greater improvements. This digital health program could provide a new tool to improve health outcomes in people with NAFLD.

**Trial Registration:** Clinicaltrials.gov NCT05426382; https://clinicaltrials.gov/study/NCT05426382

**KEYWORDS**
digital health program; nonalcoholic fatty liver disease; NAFLD; cardiometabolic health; digital therapeutics; liver; chronic hepatic; cardiometabolic; cardiovascular; cardiology; weight; acceptability; digital health; metabolic syndrome; diabetic; diabetes; diabetics; type 2; BMI; lifestyle; exercise; physical activity; coaching; diet; dietary; nutrition; nutritional; patient education; coach; feasibility; fat; body composition

**Introduction**
Nonalcoholic fatty liver disease (NAFLD) is the most common chronic liver disease in the world [1]. NAFLD is defined as >5% fat in the liver (steatosis) among people who drink moderate amounts or no alcohol and have no other chronic liver diseases [2]. NAFLD reflects a spectrum of liver pathologies, ranging from simple steatosis to a more severe condition called nonalcoholic steatohepatitis (NASH), which includes inflammation and potential scarring of the liver [3]. The major comorbidities associated with NAFLD are central obesity, type 2 diabetes mellitus (DM), dyslipidemia, and metabolic syndrome [4]. The global prevalence of NAFLD is estimated to be 25% in the general population and the rising prevalence of NAFLD parallels that of obesity and type 2 DM, since NAFLD is a comorbidity in an estimated 55% of people with type 2 DM and in up to 80% of people with obesity [4-6]. Studies have shown that around 20% to 30% of people with NAFLD progress to NASH, with its consequent risks of liver scarring, cirrhosis, end-stage liver disease, and hepatocellular carcinoma [7]. Furthermore, NAFLD and NASH are associated with cardiovascular diseases, type 2 DM, and chronic kidney disease and pose a large burden on health care systems [8-10].

Growing evidence supports a common pathophysiological mechanism between metabolic syndrome and NAFLD and NASH, which often involves insulin resistance and dysfunctional adipose tissue [11]. Currently, no pharmacological treatment is approved for NAFLD or NASH, and, according to treatment guidelines, first-line therapy should focus on lifestyle improvement with the aim of 5% to 10% weight loss [12,13]. However, reaching these goals is often difficult, and there is a need to continue exploring optimal treatment modalities for individuals with NAFLD or NASH [14].

Sidekick Health, an Icelandic digital therapeutic company, has developed a digital health program (Sidekick-241 or SK-241) specifically designed for people with metabolic conditions and NAFLD. The 12-week program is delivered through a mobile app and aims to improve lifestyle and health outcomes by focusing on improving diet, increasing activity levels, and reducing stress through behavior change. In this prospective study, we evaluated the feasibility and potential clinical impact of the 12-week digital health program on liver and cardiometabolic health in individuals with metabolic conditions and NAFLD.

**Methods**

**Trial Design**
This was an open-label, single-arm, prospective study conducted between June and September 2022 in Iceland. The study included a 12-week digital health program delivered through the Sidekick app. Screening and prepogram and postprogram clinical assessments were carried out at the Icelandic Heart Association.

**Participants**
In total, 38 individuals aged between 18 and 80 years from an ongoing population-based cohort study (The REFINE-Reykjavik Study) at The Icelandic Heart Association and individuals followed at an endocrine outpatient clinic (the Reykjavik Heart Center) were invited to participate in the study [15]. People with at least one of the following risk factors were invited to participate: BMI >30, metabolic syndrome, or type 2 DM. Individuals with type 2 DM were only included if they were on a stable dose of antidiabetes medication for the last 90 days before screening. Eligible individuals had to have the capacity to give informed consent, understand verbal and written Icelandic, own and know how to operate a smartphone, and be willing and able to comply with the study program, all scheduled visits, and procedures.

The exclusion criteria were as follows: insulin use; known or self-reported cirrhosis; alcohol consumption over 14 units/week for men or over 7 units/week for women; self-reported hepatitis B, hepatitis C, human immunodeficiency virus, or autoimmune hepatitis; vitamin E intake of >400 IU/day unless stable for 12 weeks prior to baseline; taking medications associated with liver steatosis, such as steroids, methotrexate, tamoxifen, amiodarone, tetracycline, or valproic acid; self-reported pregnancy; participation in a weight loss program; or history of or any existing medical condition (eg, ongoing cancer treatment, severe cardiopulmonary or musculoskeletal disease); magnetic resonance imaging contraindications (eg, pacemakers, aneurysm clips), or stroke or myocardial infarction in the last 6 months that, in the opinion of the primary investigator, would interfere with evaluation of the study or affect the interpretation of the results of the study.

After obtaining informed consent, participants were screened for eligibility by study staff at the Icelandic Heart Association.
Screening for NAFLD

Individuals were screened to assess if they had liver steatosis with a noninvasive ultrasonography-based controlled attenuation parameter (CAP) assessment through a FibroScan device [16]. To avoid overestimation of steatosis, we used 2 probe sizes (medium and extra-large). Individuals who met the full inclusion and exclusion criteria and had a CAP score of >294 dB/m, which represents a high likelihood of >5% liver steatosis, were eligible for participation in the study [17]. Additionally, liver stiffness measurement (LSM) was performed with vibration-controlled transient elastography (VCTE) at screening and at the 12-week follow-up visit. Individuals with an LSM score >9.7 kPa, which represents moderate-to-severe liver fibrosis (grade F3-F4), were referred to a specialist for further evaluation [18]. All individuals with a CAP score >294 dB/m had a magnetic resonance imaging proton density fat fraction (MRI-PDFF) measurement at screening and at the 12-week follow-up visit. MRI-PDFF is considered an emerging biomarker for non-invasive hepatic steatosis assessment as it is accurate, precise, quantitative, and reproducible [19].

The Digital Health Program

The SK-241 digital health program was developed by a multidisciplinary group of experts, including a clinical psychologist, nutritionist, behavioral scientists, medical doctors, and nurses at Sidekick Health. The primary focus of the program was to reduce participants’ daily dietary carbohydrate consumption and improve their overall nutrition quality in small, achievable, and sustainable steps (eg, reducing added sugars and processed foods, prioritizing protein, and increasing vegetable consumption). A secondary focus was to increase daily activity levels, improve sleep quality and reduce stress. The user interface with example screenshots from the program is shown in Figure 1.

The program included short daily missions (defined as in-app tasks for the participant to complete) aimed at increasing knowledge about NAFLD and NASH and its contributing factors and improving participants’ lifestyles for better metabolic health. The daily missions included watching short educational videos, reading brief informational content, logging meals and beverages by taking a photo of the meal, assessing on a sliding scale how healthy the meal was, and evaluating hunger and satiety before and after the meal. Other missions involved practicing mindfulness and meditation and logging daily energy levels, stress, and sleep quality. The app also provided participants with in-app health coach support (by a live person, not artificial intelligence), which provided weekly feedback on food logs and other in-app activities and opportunities for participants to ask questions as needed. The coaching element was active throughout the 12-week intervention period. Further details of the in-app content and missions are presented in Table 1 and Table S1 in Multimedia Appendix 1.

Figure 1. Example screens of the Sidekick app and the Sidekick-241 NAFLD program user interface. NAFLD: non-alcoholic fatty liver disease.
Table 1. The Sidekick-241 program content and descriptions of main missions.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food journal</td>
<td>Participants were asked to log their meals at least 3 times per week each week. During week 2, individualized goals for gradually reducing carbohydrate intake throughout the program were set based on week 1 consumption.</td>
</tr>
<tr>
<td>Step counter</td>
<td>Participants could manually log their steps each day. Individualized goals for increasing steps were set for week 2 based on week 1 step counts.</td>
</tr>
<tr>
<td>QoL PROs</td>
<td>Participants were prompted to log these measures 2 days per week on a 10-point visual-analog sliding scale.</td>
</tr>
<tr>
<td>Surveys</td>
<td>Questions about motivation levels, knowledge and attitudes relating to nutrition and physical activities were administered during weeks 1 and 2 and again during weeks 11 and 12. Questions about current food-related behaviors and potential NAFLD- or NASH-related symptoms were administered every 2 weeks.</td>
</tr>
<tr>
<td>Mindfulness</td>
<td>Participants were prompted to complete short mindfulness exercises regularly throughout the program and practice meditation 2 times per week from week 3 onwards.</td>
</tr>
<tr>
<td>Coaching</td>
<td>Feedback on weekly in-app activities was provided to participants, particularly on food logs and answers to the in-app surveys. Throughout the program, participants were also able to ask questions as needed and the coach would answer within 24 hours (weekends exempt).</td>
</tr>
</tbody>
</table>

aQoL: quality of life.
bPROs: patient-reported outcomes.

During the baseline visit, study staff assisted participants with downloading and installing the Sidekick app with the SK-241 program. A short web-based interview with the program’s health coach was offered to all participants during the first 2-3 weeks of the study to establish coach connection and accountability and to provide participants with an opportunity to ask questions. During the interview, the primary goals, main concepts, and the program’s approach to diet and weight management were explained. In addition, participants’ strengths and potential barriers to participation were discussed.

Outcome Measures and Covariates

Primary outcomes were the program’s feasibility and acceptability, as assessed by participant retention, engagement, and satisfaction after the 12-week study period. An active participant was defined as one completing at least 1 in-app mission or interacting with the health coach at least once per week. Retention was measured as the number of participants completing the 12-week program, which was defined as being active 9 of 12 weeks. Engagement was measured as the number of participants who were active during the whole 12-week period. Satisfaction with the program was assessed after program completion with the validated mHealth App Usability Questionnaire (MAUQ), which consists of 18 items and has a possible score of 0–7, with 7 being the highest potential score. The scoring can further be divided into 3 subscales reflecting ease of use (5 items), interface and satisfaction (7 items), and usefulness (6 items) [20]. In addition, detailed participant engagement with specific program features was analyzed.

Secondary outcomes were the program’s preliminary and potential clinical impact, as measured by weight loss, changes in liver fat, body composition, serum biomarkers, and other cardiometabolic risk factors (eg, blood pressure, waist and hip circumference, and step counts). Participants were assessed at baseline and at a 12-week follow-up visit for demographic information, anthropometric measures, medical history, medications, and adverse events. Liver fat content was measured and quantified at baseline and at 12 weeks using MRI-PDFF with a multiecho chemical shift-encoded gradient-echo sequence [21]. Body composition was assessed at baseline and at 12 weeks with a dual-energy X-ray absorptiometry [22]. Blood pressure was measured using an automatic blood pressure monitor. Blood samples were drawn at baseline and at the 12-week follow-up to measure complete blood count, alanine aminotransferase, aspartate aminotransferase, hemoglobin A1c (HbA1c), fasting glucose and insulin for the homeostatic model assessment of insulin resistance (HOMA-IR), total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, and high-sensitivity C-reactive protein.

Participants were administered the following questionnaires via an electronic patient-reported outcome (PRO) system at baseline and at 12 weeks: the Depression, Anxiety and Stress Scale (DASS-21), the EuroQol-5 Dimension – 5-Level (EQ-5D-5L) index, and the 8-item Morisky Medication Adherence Scale (MMAS-8) [23-25].

For exploratory outcome analysis, study participants were divided into 2 groups depending on how engaged they were with the digital health program. Those using the app 5 or more days per week were defined as highly engaged compared with those using the app less than 5 days per week, and clinical outcomes were compared to assess a potential dose-response relationship.

Statistical Analysis

As this is a feasibility study, a formal sample size calculation was not performed. The researchers aimed for 30–40 participants as this was considered a sufficiently sized sample to obtain information on practical aspects of participants’ recruitment, in-app engagement, retention, and rates of acceptance.

Changes in clinical assessments and PROs from baseline to postprogram were calculated as the mean and SD for approximately normally distributed variables (normality was analyzed with the Shapiro-Wilk test) or as the median and IQR for variables that did not satisfy normality criteria. Categorical data were calculated as frequencies and percentages. To compare baseline and postprogram outcomes, paired t tests were computed for approximately normally distributed data. In case
the normality assumption was not met, nonparametric tests were computed (Wilcoxon signed-rank tests). Unless otherwise specified, all statistical tests were performed at the 5% (2-sided) significance level. Statistical analysis was performed in Stata (StataCorp) and R (version 4.0.3; R Foundation for Statistical Computing).

All enrolled participants were included in the full analysis set. Missing data were imputed using the last observation carried forward provided that the participant was enrolled in the study and at least one of two measurements (baseline or follow-up) was collected. Moreover, missing baseline measurements in waist circumference, hip circumference, and low-density lipoprotein cholesterol were imputed for 1 participant using the next observation carried backward. The complete case analysis set included participants who attended both the baseline visit and the 12-week follow-up visit.

Ethical Considerations

This study was approved by the National Bioethics Committee of Iceland and the Data Protection Authority (22-075-VI). All participants provided informed consent before being enrolled in the study. All data was deidentified and analyzed in accordance with institutional protocols. Participants were given the option of seeking reimbursement for travel expenses not exceeding US $150 in total; no other compensation was provided. The study was registered at ClinicalTrials.gov under the trial identifier NCT05426382.

Results

Participant Characteristics

After screening and enrollment, 38 individuals were eligible to participate in the study (Figure 2). The median age of the participants was 59.5 (IQR 46.3-68.8) years, 23 (61%) were women and all were White (Table 2). Of the 38 participants, 17 (45%) had a university degree, none smoked, 34 (90%) had obesity (BMI >30), 19 (50%) had type 2 DM, 27 (71%) had hypertension, 15 (40%) had hypercholesterolemia, and 11 (29%) had a history of cardiovascular disease. Other common comorbidities included hypothyroidism (n=11, 29%), polycystic ovary syndrome (n=4, 11%), and gout (n=2, 5%). In total, 45% (n=17) of participants reported taking antidiabetic medication, 79% (n=30) antihypertensive medication, 37% (n=14) antilipidemic medication, and 37% (n=14) hypothyroid medication. Additionally, 74% (n=28) reported taking other medications, such as proton-pump inhibitors (n=11, 29%), anticoagulants (n=11, 29%), antidepressants (n=8, 21%), vitamin B12 (n=5, 13%), nonsteroidal anti-inflammatory medication (n=4, 11%), and antihistamines (n=4, 11%). During the 12-week study period, 5 (13%) participants reported medication changes: 3 (8%) started new medications (one received antibiotics, one received calcium channel blockers, and one vitamin B12) and 2 (5%) reported dosage adjustments (one for diabetes medications and one for beta blockers and antidepressants).

Figure 2. Flowchart of study participants. MRI: magnetic resonance imaging; SK: Sidekick.
Table 2. Baseline characteristics of the study participants.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants (n=38)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>23 (61)</td>
</tr>
<tr>
<td>Men</td>
<td>15 (39)</td>
</tr>
<tr>
<td><strong>Age (years), median (IQR)</strong></td>
<td>59.5 (46.3-68.8)</td>
</tr>
<tr>
<td><strong>Ethnicity: White, n (%)</strong></td>
<td>38 (100)</td>
</tr>
<tr>
<td><strong>Work status, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>18 (47)</td>
</tr>
<tr>
<td>Part-time</td>
<td>5 (13)</td>
</tr>
<tr>
<td>Not in labor market</td>
<td>15 (39)</td>
</tr>
<tr>
<td>Pension</td>
<td>11 (29)</td>
</tr>
<tr>
<td>Disability</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Sick leave</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1 (3)</td>
</tr>
<tr>
<td><strong>Educational level, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>University degree</td>
<td>17 (45)</td>
</tr>
<tr>
<td>Trades or vocational school or equivalent</td>
<td>12 (32)</td>
</tr>
<tr>
<td>Primary education or less</td>
<td>6 (16)</td>
</tr>
<tr>
<td>Secondary or matriculate</td>
<td>3 (8)</td>
</tr>
<tr>
<td><strong>Smoking status, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Never smoked</td>
<td>20 (53)</td>
</tr>
<tr>
<td>Former smoker</td>
<td>18 (47)</td>
</tr>
<tr>
<td><strong>Comorbidities, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>19 (50)</td>
</tr>
<tr>
<td>BMI &gt;30</td>
<td>34 (89)</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>15 (39)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>27 (71)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>11 (29)</td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>11 (29)</td>
</tr>
<tr>
<td>Polycystic ovary disease</td>
<td>4 (11)</td>
</tr>
<tr>
<td>Gout</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Other</td>
<td>23 (61)</td>
</tr>
</tbody>
</table>

**Retention and Engagement in the 12-Week Program**

Of the 38 participants, 34 (89%) completed the 12-week program, 29 (76%) were engaged during the whole study period, and 22 (58%) were highly engaged (defined as visiting the app at least 5 days per week) (Table 3). Engagement and retention in the app were similar between those younger or older than 60 years and between men and women (data not shown). Participants were active in-app on a median of 81 (IQR 45.8-84.0) of 84 days or 6.8 (IQR 4.6-7.0) days per week on average and completed an average of 6.9 (SD 2.9) daily missions. Over the course of the study, the health coach sent an average 23.5 (SD 10.3) messages to participants, while participants sent and average of 15.5 (SD 12.4) messages to the coach, who responded within 1.2 (SD 0.9) days. The median MAUQ score was 6.3 (IQR 5.8-6.7) of 7, suggesting high satisfaction with the program among participants.
Table 3. Overall retention, engagement, and satisfaction.

<table>
<thead>
<tr>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary endpoints</strong></td>
<td></td>
</tr>
<tr>
<td>Retention(^a), n (%)</td>
<td>34 (89)</td>
</tr>
<tr>
<td>Engagement(^b), n (%)</td>
<td>29 (76)</td>
</tr>
<tr>
<td><strong>Satisfaction</strong> median (IQR)</td>
<td></td>
</tr>
<tr>
<td>MAUQ(^d) total score</td>
<td>6.3 (5.8-6.7)</td>
</tr>
<tr>
<td>Ease of use (mean of MAUQ items 1 to 5)</td>
<td>6.4 (5.6-6.8)</td>
</tr>
<tr>
<td>Interface and satisfaction (mean of MAUQ items 6 to 12)</td>
<td>6.3 (5.9-6.9)</td>
</tr>
<tr>
<td>Usefulness (mean of MAUQ items 13 to 18)</td>
<td>6.0 (5.5-6.7)</td>
</tr>
<tr>
<td><strong>Exploratory engagement metrics</strong></td>
<td></td>
</tr>
<tr>
<td>Average active(^e) days per week (0-7), median (IQR)</td>
<td>6.8 (4.6-7.0)</td>
</tr>
<tr>
<td>Average total active days (0-84), median (IQR)</td>
<td>81 (45.8-84.0)</td>
</tr>
<tr>
<td>Average daily missions completed(^f), mean (SD)</td>
<td>6.9 (2.9)</td>
</tr>
<tr>
<td>Average daily missions assigned, mean (SD)</td>
<td>5.7 (0.49)</td>
</tr>
<tr>
<td>Participants who were active &gt;5 days every week, n (%)</td>
<td>22 (58)</td>
</tr>
<tr>
<td>Average number of messages sent by participants, mean (SD)</td>
<td>15.5 (12.4)</td>
</tr>
<tr>
<td>Average number of messages received by participants, mean (SD)</td>
<td>23.5 (10.3)</td>
</tr>
</tbody>
</table>

\(^a\)Retention was defined as participants who completed the program, being active for 9 of 12 weeks. Being active was defined as completing at least 1 in-app mission or interacting at least once in that week with the health coach.

\(^b\)Engagement was defined as participants who were active for the entire study period.

\(^c\)Satisfaction was measured using the MAUQ.

\(^d\)MAUQ: mHealth App Usability Questionnaire.

\(^e\)An active day was defined as a day in which the participant completed at least 1 in-app mission or interacted with the coach.

\(^f\)The participants receive daily assigned missions but also had the opportunity to complete additional missions within the app, thereby surpassing the number of assigned missions.

**Metabolic Parameters**

The mean weight loss was 3.5 (SD 3.7) kg (\(P<.001\)), or 3.2% (SD 3.4%) (Table 4). The median body fat percentage changed from 46.6% (IQR 39.4%-52.4%) to 44.3% (IQR 37.8%-52.2%) (\(P<.001\)) and the mean fat mass from 50.3 (SD 13.8) kg to 48.1 (SD 14.5) kg (\(P<.001\)). These improvements in body composition were accompanied by reduced MRI-PDFF liver fat values: in the full analysis set (n=38), the mean liver fat percentage significantly decreased from 12.3% (SD 7.1%) to 10.1% (SD 6.5%; \(P<.001\)), representing a mean relative change of 19.4% (SD 23.9%) (Table 4). In the complete case analysis set (n=34), mean liver fat was reduced from 12.4% (SD 6.9%) to 9.9% (SD 6.3%; \(P<.001\)) with a corresponding mean relative change of 21.6% (SD 24.2%).
Table 4. Differences in anthropometric, biochemical, and clinical measurements at baseline and after 12 weeks for the full analysis set (n=38).

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Week 12</th>
<th>Change from baseline to week 12</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anthropometry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg), mean (SD)</td>
<td>110.0 (18.5)</td>
<td>106.5 (18.4)</td>
<td>3.5 (3.7)</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Relative percentage weight change&lt;sup&gt;b&lt;/sup&gt;, mean (SD)</td>
<td>N/A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>N/A</td>
<td>3.2 (3.4)</td>
<td>N/A</td>
</tr>
<tr>
<td>BMI (kg/m&lt;sup&gt;2&lt;/sup&gt;), mean (SD)</td>
<td>37.6 (5.8)</td>
<td>36.4 (5.8)</td>
<td>1.2 (1.3)</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Waist circumference (cm), mean (SD)</td>
<td>123.8 (12.2)</td>
<td>119.9 (12.2)</td>
<td>4.0 (5.1)</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hip circumference (cm), mean (SD)</td>
<td>125.1 (14.0)</td>
<td>123.2 (13.3)</td>
<td>1.8 (0.0 to 4.9)</td>
<td>.01&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Waist to hip ratio, median (IQR)</td>
<td>1.00 (0.95 to 1.03)</td>
<td>0.99 (0.92 to 1.03)</td>
<td>0.00 (–0.01 to 0.03)</td>
<td>.09&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Liver assessment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver fat MRI-PDFF&lt;sup&gt;e&lt;/sup&gt; (%)</td>
<td>12.3 (7.1)</td>
<td>10.1 (6.5)</td>
<td>2.2 (2.9)</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Liver fat MRI-PDFF relative change&lt;sup&gt;b&lt;/sup&gt; (%)</td>
<td>N/A</td>
<td>N/A</td>
<td>19.4 (23.9)</td>
<td>N/A</td>
</tr>
<tr>
<td>Liver stiffness measure (kPa), median (IQR)</td>
<td>6.4 (5.2 to 9.6)</td>
<td>6.6 (5.3 to 8.4)</td>
<td>0.2 (–0.3 to 1.6)</td>
<td>.11&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>CAP&lt;sup&gt;f&lt;/sup&gt; score (dB/m), mean (SD)</td>
<td>343.6 (34.8)</td>
<td>310.3 (47.2)</td>
<td>33.3 (39.7)</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Body composition&lt;sup&gt;g&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total body region fat (%), median (IQR)</td>
<td>46.6 (39.4 to 52.4)</td>
<td>44.3 (37.8 to 52.2)</td>
<td>0.9 (1.4)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fat mass (kg), mean (SD)</td>
<td>50.3 (13.8)</td>
<td>48.1 (14.5)</td>
<td>2.2 (2.7)</td>
<td>&lt;.001&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Lean mass (kg), mean (SD)</td>
<td>56.3 (10.1)</td>
<td>55.6 (9.7)</td>
<td>0.7 (1.7)</td>
<td>.008&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Blood pressure (mmHg), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>141.4 (17.1)</td>
<td>135.4 (17.3)</td>
<td>6.0 (13.5)</td>
<td>0.009&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Diastolic</td>
<td>83.6 (7.4)</td>
<td>82.5 (7.4)</td>
<td>1.2 (7.7)</td>
<td>.36&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Biochemical measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HbA&lt;sub&gt;1c&lt;/sub&gt;&lt;sup&gt;h&lt;/sup&gt; (mmol/mol), median (IQR)</td>
<td>60.0 (56.0 to 66.8)</td>
<td>60.0 (54.3 to 64.0)</td>
<td>0.5 (–0.7 to 3.8)</td>
<td>.03&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>S-glucose&lt;sup&gt;i&lt;/sup&gt; (mmol/L), median (IQR)</td>
<td>6.2 (5.3 to 7.4)</td>
<td>6.3 (5.4 to 6.9)</td>
<td>0.0 (–0.3 to 0.4)</td>
<td>.64&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>S-insulin&lt;sup&gt;j&lt;/sup&gt; (µU/ml), median (IQR)</td>
<td>21.1 (16.4 to 27.9)</td>
<td>19.0 (13.0 to 25.0)</td>
<td>3.2 (0.0 to 5.4)</td>
<td>.003&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>HOMA-IR&lt;sup&gt;k&lt;/sup&gt; (mmol/L), median (IQR)</td>
<td>5.8 (4.3 to 8.4)</td>
<td>4.8 (3.6 to 7.2)</td>
<td>0.4 (–0.2 to 2.1)</td>
<td>.02&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Total cholesterol (mmol/L), mean (SD)</td>
<td>4.9 (1.3)</td>
<td>4.8 (1.2)</td>
<td>0.0 (–0.2 to 0.2)</td>
<td>&gt;.99&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>LDL-C&lt;sup;l&lt;/sup&gt; (mmol/L), mean (SD)</td>
<td>2.9 (1.1)</td>
<td>2.9 (1.1)</td>
<td>–0.1 (–0.3 to 0.1)</td>
<td>.18&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>HDL-C&lt;sup&gt;m&lt;/sup&gt; (mmol/L), mean (SD)</td>
<td>1.11 (0.23)</td>
<td>1.12 (0.19)</td>
<td>–0.01 (0.12)</td>
<td>.56&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Triglycerides (mmol/L), median (IQR)</td>
<td>1.88 (1.35 to 2.45)</td>
<td>1.68 (1.21 to 1.90)</td>
<td>0.14 (0.00 to 0.47)</td>
<td>.003&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>hs-CRP&lt;sup&gt;n&lt;/sup&gt; (mg/L), median (IQR)</td>
<td>3.0 (1.2 to 5.2)</td>
<td>2.5 (1.1 to 3.9)</td>
<td>0.1 (–0.1 to 0.7)</td>
<td>.14&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>ALAT&lt;sup,o&lt;/sup&gt; (IU/L), median (IQR)</td>
<td>21.4 (18.2 to 30.2)</td>
<td>23.2 (18.4 to 32.0)</td>
<td>0.0 (–6.8 to 2.8)</td>
<td>.37&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>ASAT&lt;sup,p&lt;/sup&gt; (IU/L), median (IQR)</td>
<td>20.8 (17.9 to 24.8)</td>
<td>22.3 (18.0 to 25.5)</td>
<td>0.4 (–2.5 to 2.5)</td>
<td>.53&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>FIB-4 Index, median (IQR)</td>
<td>1.08 (0.78 to 1.34)</td>
<td>1.08 (0.75 to 1.21)</td>
<td>0.01 (–0.06 to 0.07)</td>
<td>.58&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Analyzed with a paired t test.
<sup>b</sup>Percentage change calculated as the average over individual relative changes.
<sup>c</sup>N/A: not applicable.
<sup>d</sup>Analyzed with a Wilcoxon signed-rank test.
During the study, the distribution of steatosis levels changed. At baseline, the 10%-15% liver steatosis category had the highest frequency with 32% (n=12) of participants. At follow-up, the 5%-10% liver steatosis category had the highest frequency with 34% (n=13) of participants (Table 5). We additionally found a significant correlation between weight loss and absolute ($r=0.48$, $P=.004$) and relative ($r=0.72$, $P<.001$) liver fat changes measured by MRI-PDFF.

According to the FIB-4 data, 4 of the 38 participants were classified as having a high risk of fibrosis at baseline, of which 1 individual regressed to intermediate risk at the 12-week follow-up visit. Most participants (n=27, 71%) had a low risk of fibrosis at baseline according to the FIB-4. At the 12-week follow-up, this percentage had gone up to 79% (n=30).

Mean systolic blood pressure significantly decreased by 6.0 (SD 13.5) mmHg ($P=.009$), and this was not explained by changes in medication or medication adherence (Table 4). There was no significant difference in diastolic blood pressure.

Participants recorded on average 3085 (SD 2246) daily steps in the first week and 4664 (SD 3780) daily steps in the last week, representing a significant increase of 1579 steps per day ($P=.02$).

While participants’ average baseline fasting insulin and HOMA-IR levels indicated insulin resistance, we found a significant decrease in serum insulin levels (median 3.2, IQR 0.0-5.4 µU/ml; $P=.003$), HOMA-IR levels (median 0.4, IQR -0.2 to 2.1 mmol/L; $P=.02$), and HbA1c levels (median 0.5, IQR -0.7 to 3.8 mmol/mol; $P=.03$) (Table 4), suggesting improved glycemic control. In addition, triglyceride levels significantly decreased by a median of 0.14 (IQR 0.00-0.47) mmol/L ($P=.003$), and median high-sensitivity C-reactive protein levels decreased from 3.0 (IQR 1.2-5.2) mg/L to 2.5 (IQR 1.1-3.9) mg/L ($P=.14$) (Table 4), representing improvements in those cardiovascular risk factors.

We did not find any significant change in cholesterol levels, nor any significant changes in PRO scores of health-related quality of life, mental health, or medication adherence from preprogram to postprogram (Table 6).

Table 5. Distributions of liver fat percentage categories based on magnetic resonance imaging proton density fat fraction liver fat values at baseline and at the 12-week follow-up (n=38). All participants with >5% liver fat at baseline had stage 1 steatosis according to the standardized Nonalcoholic Steatohepatitis Clinical Research Network histologic scoring system for nonalcoholic fatty liver disease [26].

<table>
<thead>
<tr>
<th>Liver fat category (%)</th>
<th>Participants at baseline, n</th>
<th>Participants at week 12, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>5-10</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>10-15</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>15-20</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>20-25</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>25-30</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

*Limits for the presented ranges correspond to values greater than or equal to for lower limits and less than for upper limits.
Table 6. Differences in patient reported outcomes (PROs) at baseline and after 12 weeks for the full analysis set (n=38).

<table>
<thead>
<tr>
<th>PROs</th>
<th>Baseline</th>
<th>Week 12</th>
<th>Change from baseline to week 12</th>
<th>P valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ-5D-5L indexb</td>
<td>0.8 (0.7 to 0.9)</td>
<td>0.9 (0.8 to 1.0)</td>
<td>0.0 (~0.1 to 0.0)</td>
<td>.36</td>
</tr>
<tr>
<td>DASS-21c, median (IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score (0 to 56)</td>
<td>5.5 (2.0 to 13.0)</td>
<td>6.0 (2.0 to 11.0)</td>
<td>0.0 (~2.0 to 2.0)</td>
<td>.95</td>
</tr>
<tr>
<td>Depression score</td>
<td>1.0 (0.3 to 4.5)</td>
<td>1.0 (0.0 to 3.0)</td>
<td>0.0 (0.0 to 1.0)</td>
<td>—</td>
</tr>
<tr>
<td>Anxiety score</td>
<td>1.0 (0.0 to 2.0)</td>
<td>1.0 (0.0 to 2.0)</td>
<td>0.0 (~0.7 to 1.0)</td>
<td>—</td>
</tr>
<tr>
<td>Stress score</td>
<td>3.0 (1.0 to 6.0)</td>
<td>3.0 (0.3 to 6.0)</td>
<td>0.0 (~1.0 to 1.0)</td>
<td>—</td>
</tr>
<tr>
<td>MMAS-8d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score (0 to 8)</td>
<td>7.0 (6.8 to 8.0)</td>
<td>7.0 (6.8 to 8.0)</td>
<td>0.0 (~0.7 to 0.0)</td>
<td>.68</td>
</tr>
<tr>
<td>High adherence (=8), n (%)</td>
<td>14 (37)</td>
<td>15 (39)</td>
<td>N/Ae</td>
<td>—</td>
</tr>
<tr>
<td>Moderate adherence (6 to 7), n (%)</td>
<td>15 (39)</td>
<td>15 (39)</td>
<td>N/A</td>
<td>—</td>
</tr>
<tr>
<td>Low adherence (&lt;6), n (%)</td>
<td>9 (24)</td>
<td>8 (21)</td>
<td>N/A</td>
<td>—</td>
</tr>
</tbody>
</table>

a Analyzed with Wilcoxon signed-rank test.
bEQ-5D-5L index: EuroQol-5 Dimension – 5-Level index.
cDASS-21: Depression, Anxiety, and Stress Scale - 21 Items.
dNot available.
eMMAS-8: 8-item Morisky Medication Adherence Scale.
fN/A: not applicable.

**Associations Between App Engagement and Clinical Outcomes**

An exploratory analysis was performed to assess the relationship between participants’ in-app activity and their clinical outcomes. We found that participants who were highly engaged (visited the app at least 5 days per week) had greater weight loss and liver fat reduction (Table S2 in Multimedia Appendix 1) compared with those who were less engaged. In a complete case analysis, participants who were highly engaged (n=22) lost on average 5.1 (SD 3.8) kg and achieved a 27.5% relative reduction in liver fat, while those who were active on fewer than 5 days a week (n=12) lost on average 1.8 (SD 2.2) kg and achieved 10.8% relative reduction in liver fat. Moreover, highly engaged participants were significantly more likely to achieve a relative weight loss of at least 3% (P=.001) or 5% (P=.02) compared with those who were less engaged (Fisher exact tests). Taken together, these results suggest that higher engagement with the digital program may be associated with improved metabolic health.

**Adverse Events**

In total, 9 adverse events were reported, all of which were of mild to moderate intensity with no serious adverse events (Table S3 in Multimedia Appendix 1). No adverse events were considered to have a causal relationship to the digital health program, as assessed by the primary investigator.

**Discussion**

**Principal Findings**

This study demonstrated that the 12-week-long digital health program, SK-241, was feasible given its high retention, engagement, and satisfaction among people with NAFLD. Cardiometabolic health and liver-specific outcomes improved over the 12-week study period with a significant weight loss and reductions in fat mass, liver fat, systolic blood pressure, triglycerides, insulin, and HbA1c levels.

Digital behavioral programs can be effective at targeting weight loss among people with chronic conditions [27]. Increasing evidence shows that programs—whether digital or face-to-face—with a holistic approach can also be effective for people with NAFLD, where weight loss is a major component of disease management. A recent randomized controlled study from Singapore including 108 adults with NAFLD randomized either to lifestyle advice by a trained nurse or using a lifestyle mobile app in addition to receiving advice by a dietitian showed that the mobile app group had a 5-fold higher likelihood of achieving ≥5% weight loss compared with the control group at 6 months [28].

Previous studies suggest that digital solutions can be as effective as face-to-face behavioral change programs, but engagement with the digital program is an important component of efficacy [29]. Indeed, an important finding of this study was the correlation between participants’ in-app engagement and their clinical outcomes; this shows that maintaining engagement and interest is key to reaching the desired clinical improvements. Program engagement may be influenced by several factors, such as recruitment methods, participant characteristics, app design, and the level of support, such as coaching [30,31]. Coaching in particular may be essential to drive engagement as it encourages accountability and may increase motivation [30]. The regular contact that participants had with the coach in the SK-241 program may have contributed to the low attrition and high engagement in this study. Should larger implementation of this approach be considered, it is important to maintain high engagement levels to ensure continued benefit for participants.
intervention take place, then coaching would be an integral part, at least in the initial stages of the program.

Lifestyle interventions consisting of diet, exercise, and weight loss are recommended to individuals with NAFLD according to treatment guidelines [2]. The primary driver of NAFLD is overnutrition, which causes expansion of adipose deposits and macrophage infiltration into the visceral adipose tissue, creating a proinflammatory state that promotes insulin resistance [32,33]. The resulting imbalance in lipid metabolism leads to the formation of lipotoxic lipids that contribute to cellular stress, including oxidative stress, inflammasome activation, and apoptotic cell death [34,35]. Central obesity is also an important driver of insulin resistance and proinflammatory signaling [36].

In this 12-week study, the mean waist circumference was significantly reduced by 4.0 cm, and body weight by 3.2% on average; these are encouraging results, considering that a 3%-5% weight loss can lead to a reduction in hepatic steatosis [37]. In addition, we found a correlation between weight loss and MRI-PDFF liver fat fraction changes. This is in line with a previous report of greater weight loss leading to more significant improvements in liver histopathology, and studies have shown that a ≥30% relative decline in liver fat by MRI-PDFF is associated with histopathological improvements in NASH [38-40]. Participants in this study were able to decrease their waist circumference and body weight and had an average of around a 20% relative reduction in liver fat by MRI-PDFF, with a subset of participants achieving a 30% relative decline, which might lead to improved NAFLD and NASH histopathology.

Despite the risk of progressive liver disease, the leading cause of death in people with NAFLD is cardiovascular disease [10]. This is likely due to risk factors that are shared between NAFLD and cardiovascular diseases, although it is unclear to what extent NAFLD has a direct causative role in the development of cardiovascular disease [41]. Therefore, it was important to see significant improvement in cardiovascular risk factors in our study, such as a decrease in systolic blood pressure, triglycerides, insulin, and HbA1c. The increased physical activity in our study as measured by the in-app step counter and the correlation between in-app activity and weight loss suggest that the digital program may successfully engage participants in behaviors that lead to more weight loss, which in turn may hypothetically improve liver function and glycemic control. Regular tracking of meals and physical activity and completing the in-app PROs may help people become more aware of their habits, while the education and the coach’s feedback and support may give them the necessary tools to change their behaviors. We did not find any significant changes in the PRO scores of health-related quality of life and mental health, which was most likely due to the short duration of the study and the small size of the cohort.

Furthermore, studies have shown that health care utilization and expenditure are particularly high among people with NAFLD and NASH [42,43]. Therefore, there is a great need for early identification and effective management of people with NAFLD to minimize the comorbidity burden and health care costs.

The fibrosis risk among study participants was assessed both with a VCTE FibroScan LSM and by calculating the FIB-4 index score from participants’ age and the serum alanine aminotransferase, aspartate aminotransferase, and platelet count. The results indicated that a few participants had an intermediate to high risk of having liver fibrosis (data not shown) and could be referred to as probable patients with NASH, thereby suggesting that the digital health program might be feasible for individuals with NASH in addition to those with NAFLD. However, both of these measurements have their limitations and need to be interpreted cautiously. VCTE can rule out advanced fibrosis but often leads to false positive results in NAFLD, while the FIB-4 score might overestimate fibrosis in populations older than 65 years and is considered to have a low positive predictive value for identifying advanced fibrosis [44,45].

Strengths and Limitations

A strength of this study was the high engagement and completion rate, as these are well known issues of digital health programs [46]. In addition, the holistic nature of the program, developed by a multidisciplinary team of experts and focused on multiple aspects of participants’ lifestyle, combined with the regular support provided by the coach, can be considered a strength. A further strength was the length of the program, which allowed sufficient time to assess meaningful changes in engagement and clinical outcomes.

Limitations of this study included the single-arm design, which limits the interpretation and generalizability of our findings. The lack of a control group made it difficult to directly infer the clinical benefit of digital program, thus the secondary outcomes relating to clinical efficacy should be interpreted with caution. The observed clinical improvements should also be interpreted in context with the short duration of the health program, as sustaining improvements can be challenging after short-term behavioral interventions. It should also be acknowledged that a seasonal increase in activity levels may have contributed to the observed changes, as the study began in early summer when people tend to be more physically active. Furthermore, all the participants were White and around 50% had a relatively high education level. Higher education level has been associated with a lower burden of traditional cardiovascular risk factors [47]. Previous studies have shown an association between socioeconomic status and NAFLD, where poverty seems to be a risk factor for developing NAFLD independent of other known risk factors, such as type 2 DM and obesity, and food insecurity is associated with developing NAFLD and advanced fibrosis [48]. Education and smoking status may have affected engagement with the digital health program and, therefore, the generalizability of these results to a wider population may be limited and future trials should recruit a more diverse group of participants to assess the efficacy of the program [49].

Conclusions

The 12-week-long digital health program was feasible for individuals with NAFLD, showing high user engagement, retention, and satisfaction. Improved liver-specific and cardiometabolic health was observed and more engaged...
participants showed greater improvements. This NAFLD digital health program could provide a new tool to improve health outcomes in people with NAFLD.

Acknowledgments
Judit Mészáros and Joanna McCarter provided medical writing and editing support. We would also like to thank the staff at the Icelandic Heart Association who helped with the study procedures, as well as the people who participated in the study. The study was funded by Sidekick Health. Use of the MMAS-8 is protected by US and international copyright laws. Permission for use is required. A Licensure agreement is available from: MMAR, LLC., www.moriskyscale.com.

Data Availability
The data sets generated or analyzed during this study are not publicly available due to restrictions in the informed consent form. Additional summary statistics will be provided upon on reasonable request.

Authors' Contributions
SB, HU, EFG, KS, TG, TK, SO, and VG contributed to conceptualization. SB, HU, EFG, KS, AI, BD, GEAM, TG, GB, SS, SO, and VG contributed to the methodology. SB, HU, EFG, AI, BD, SS, SO, and VG contributed to the investigation. SB and HU wrote the original draft. SB, HU, EFG, KS, AI, BD, GEAM, TG, TK, GB, SS, SO, and VG revised and edited the manuscript. SO acquired funding. SO and VG procured resources and supervised the study.

Conflicts of Interest
HU, EFG, AI, BD, KS, TK, GEAM, and TG are employed by Sidekick Health. SO is an employee and cofounder of Sidekick Health. SB received consultancy fees from Sidekick Health during the study period. SS, GB, and VG have no competing interests to declare.

Multimedia Appendix 1
Supplementary tables.

References


Abbreviations

CAP: controlled attenuation parameter
DASS-21: Depression, Anxiety, Stress Scale, 21 Items
DM: diabetes mellitus
EQ-5D-5L index: EuroQol-5 Dimension – 5-Level index
HbA1c: hemoglobin A1c
HOMA-IR: homeostatic model assessment for insulin resistance
LSM: liver stiffness measurement
MAUQ: mHealth App Usability Questionnaire
MMAS-8: 8-item Morisky Medication Adherence Scale
MRI-PDFF: magnetic resonance imaging proton density fat fraction
NAFLD: nonalcoholic fatty liver disease
NASH: nonalcoholic steatohepatitis
ppt: percentage points
PRO: patient-reported outcome
s-insulin: Serum insulin
Total-C: total cholesterol
VCTE: vibration-controlled transient elastography

©Sigridur Björnsdottir, Hildigunnur Ulfsdottir, Elias Freyr Gudmundsson, Kolbrun Sveinsdottir, Ari Pall Isberg, Bartosz Dobies, Gudlaug Erla Akerlie Magnusdottir, Thrudur Gunnarsson, Tekla Karlsson, Gudlaug Bjornsdottir, Sigurdur Sigurdsson, Saemundur Oddsson, Vilmundur Gudnason. Originally published in JMIR Cardio (https://cardio.jmir.org), 15.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cardio, is properly cited. The complete bibliographic information, a link to the original publication on https://cardio.jmir.org, as well as this copyright and license information must be included.
Factors That Influence Patient Satisfaction With the Service Quality of Home-Based Teleconsultation During the COVID-19 Pandemic: Cross-Sectional Survey Study

Guangxia Meng¹, MSN; Carrie McAiney¹, PhD; Ian McKillop¹, PhD; Christopher M Perlman¹, PhD; Shu-Feng Tsao¹, PhD; Helen Chen¹, PhD

School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Helen Chen, PhD
School of Public Health Sciences
University of Waterloo
200 University Ave W
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567 ext 42131
Email: helen.chen@uwatertloo.ca

Abstract

Background: Ontario stroke prevention clinics primarily held in-person visits before the COVID-19 pandemic and then had to shift to a home-based teleconsultation delivery model using telephone or video to provide services during the pandemic. This change may have affected service quality and patient experiences.

Objective: This study seeks to understand patient satisfaction with Ontario stroke prevention clinics’ rapid shift to a home-based teleconsultation delivery model used during the COVID-19 pandemic. The research question explores explanatory factors affecting patient satisfaction.

Methods: Using a cross-sectional service performance model, we surveyed patients who received telephone or video consultations at 2 Ontario stroke prevention clinics in 2021. This survey included closed- and open-ended questions. We used logistic regression and qualitative content analysis to understand factors affecting patient satisfaction with the quality of home-based teleconsultation services.

Results: The overall response rate to the web survey was 37.2% (128/344). The quantitative analysis was based on 110 responses, whereas the qualitative analysis included 97 responses. Logistic regression results revealed that responsiveness (adjusted odds ratio [AOR] 0.034, 95% CI 0.006-0.188; \( P < .001 \)) and empathy (AOR 0.116, 95% CI 0.017-0.800; \( P = .03 \)) were significant factors negatively associated with low satisfaction (scores of 1, 2, or 3 out of 5). The only characteristic positively associated with low satisfaction was when survey consent was provided by the substitute decision maker (AOR 6.592, 95% CI 1.452-29.927; \( P = .02 \)). In the qualitative content analysis, patients with both low and high global satisfaction scores shared the same factors of service dissatisfaction (assurance, reliability, and empathy). The main subcategories associated with dissatisfaction were missing clinical activities, inadequate communication, administrative process issues, and absence of personal connection. Conversely, the high-satisfaction group offered more positive feedback on assurance, reliability, and empathy, as well as on having a competent clinician, appropriate patient selection, and excellent communication and empathy skills.

Conclusions: The insights gained from this study can be considered when designing home-based teleconsultation services to enhance patient experiences in stroke prevention care.

(JMIR Cardio 2024;8:e51439) doi:10.2196/51439

KEYWORDS

teleconsultation; secondary stroke prevention; telemedicine; service quality; patient satisfaction
Introduction

Secondary Stroke Prevention and Ontario Stroke Prevention Clinics

As of 2019, stroke was the second leading cause of disability worldwide for people aged >50 years, and it was the fourth leading cause of death in Canada from 2017 to 2019 [1,2]. The 36% decline in stroke mortality from 1990 to 2016 can be attributed to better prevention and management of stroke risks [3]. The INTERSTROKE study found that 90% of strokes are preventable owing to modifiable risk factors, including disease-related and behavioral lifestyle factors [4]. Secondary stroke prevention is crucial, as there is up to a 10% risk of recurrent stroke within 90 days of a transient ischemic attack (TIA) or minor stroke [5].

Approximately 80% of patients with minor stroke discharged from the emergency department in Ontario are referred to stroke prevention services [6]. Stroke prevention clinics provide rapid assessments, diagnostic tests, treatments, prevention, and education to reduce the risk of recurrent stroke [7]. Ontario’s 41 stroke prevention clinics are integral to publicly funded health systems [7]. Stroke prevention clinic services are associated with a 25% reduction in mortality [8].

Before the COVID-19 pandemic, stroke prevention care in Ontario was predominantly delivered through in-person consultations. A small percentage of rural and northern Ontario stroke prevention clinics used teleconsultation at local satellite clinics to address access challenges [9]. One stroke prevention clinic conducted a pilot project offering follow-up home video visits from August 2018 to September 2019 [10]. The video consultation produced higher patient satisfaction, was considered safe by physicians, and was shown to be cost-effective in reducing health care costs and patient expenses [10].

In this study, home-based teleconsultation was defined as a synchronous consultation between a clinical service provider and a patient in their home to provide diagnostic or therapeutic advice through telephone or videoconference [11]. Despite a handful of cases in which home-based teleconsultation was used for follow-up care, before the COVID-19 pandemic [10], Ontario stroke prevention clinics had never used home-based teleconsultation to conduct synchronous, interactive, in-home patient visits for new referrals. A survey of >3000 Canadians with stroke, heart disease, or vascular impairment conducted in the spring of 2021 showed that 80% of respondents had had a teleconsultation during the pandemic [12]. The effect of the rapid change from in-person visits to home-based teleconsultation during the COVID-19 pandemic on patients’ experiences was unknown. Patients with stroke are often older adults with multiple chronic conditions [13]. Older adults have lower telehealth service use overall [14]. The impact of service mode change on the older population of stroke prevention clinics needs further exploration.

Service Quality and Patient Satisfaction

Although the rapid transition to home-based teleconsultation may be a temporary response to the COVID-19 pandemic, it offers a significant opportunity to examine the service quality of home-based teleconsultation in stroke prevention clinics. Delivering safe, high-quality health services is the primary goal of health systems [15]. The literature’s definition of health care service quality is commonly described as including 2 aspects. One views health care service quality as characteristics and features that meet clinicians’ predetermined specifications and standards (such as professional or ethical standards); the other views it as characteristics and features that meet or exceed patients’ needs and expectations [16]. Patients often cannot accurately assess the internal service quality as they lack the medical knowledge to judge [17]; however, patient satisfaction with a medical service is the primary determinant of service quality [18]. Patient satisfaction is essential for and meaningful to delivering high-quality care [19].

Patient satisfaction is generally regarded as patients’ perception of care delivery as well as how their health needs have been addressed [20,21]. Patient satisfaction can be examined using direct and indirect indicators [21]. First, service quality can be measured by directly asking patients to rate their satisfaction with service quality via, for example, a single item with response options ranging from very dissatisfied to very satisfied [21]. However, the shortfall of single-item measurement is that we cannot evaluate or identify a specific aspect of service quality [21]. The alternative approach is to ask patients to rate their experience of different aspects of care, but this indirect measure has the weakness of preemptive assumptions about the determinants of service quality [21]. To obtain an accurate measurement of patient satisfaction, we applied direct and indirect measurements. We asked one question on global satisfaction and applied a theory-guided questionnaire to assess patient satisfaction.

Service Performance Model

As health care quality is multidimensional, we chose an appropriate service quality model (SERVQUAL) to assess patient satisfaction. Examples of existing health care SERVQUALs include the SERVQUAL [22] and its derivative service performance model (SERVPERF) [23], Total Quality Management [24], Health Quality Model [25], service quality for a public hospital [26], and hospital quality model [27]. The Health Quality Model, service quality for a public hospital, and hospital quality model are derivatives of the SERVQUAL model and assess the care quality of hospital inpatient services. The SERVQUAL and SERVPERF models include 5 dimensions: tangibles, reliability, responsiveness, assurance, and empathy [22]. SERVQUAL tends to measure the difference between one’s expectations and the actual performance of the service [22]. SERVPERF only focuses on the actual performance, and studies have shown that service quality is appropriately modeled using SERVPERF as an antecedent of satisfaction [23,28]. There are 22 service attributes listed within the 5 service dimensions of the SERVPERF model [29]. Textbox 1 presents each dimension and item in detail.
Textbox 1. Dimensions and items of the service performance model (adapted from Zeithaml et al [29]).

<table>
<thead>
<tr>
<th>Tangibles: facilities, equipment, and the presence of personnel</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Up-to-date equipment</td>
</tr>
<tr>
<td>• Visually appealing physical facilities</td>
</tr>
<tr>
<td>• Neat-appearing employees</td>
</tr>
<tr>
<td>• Visually appealing materials associated with the service</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reliability: ability to perform the promised service responsibly and accurately</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The company keeps its promises to do something by a certain time</td>
</tr>
<tr>
<td>• The company shows a sincere interest in solving the customer’s problem</td>
</tr>
<tr>
<td>• The company performs the service right the first time</td>
</tr>
<tr>
<td>• The company provides its services at the time it promises to do so</td>
</tr>
<tr>
<td>• The company insists on error-free records</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Responsiveness: willingness to provide help and a prompt service to customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Employees of the company tell customers exactly when services will be performed</td>
</tr>
<tr>
<td>• Employees of the company provide a prompt service to customers</td>
</tr>
<tr>
<td>• Employees of the company are always willing to help customers</td>
</tr>
<tr>
<td>• Employees of the company are never too busy to respond to customers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assurance: the knowledge and courtesy of employees and their ability to inspire trust and confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The behavior of employees of the company instills confidence in customers</td>
</tr>
<tr>
<td>• Customers of the company feel safe in their transactions</td>
</tr>
<tr>
<td>• Employees of the company are consistently courteous with customers</td>
</tr>
<tr>
<td>• Employees of the company have the knowledge to answer customers’ questions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Empathy: caring and understanding, which a company provides or offers its customers in terms of its individualized and personalized attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The company gives customers individual attention</td>
</tr>
<tr>
<td>• The company has operating hours convenient to all its customers</td>
</tr>
<tr>
<td>• Employees of the company give customers personal attention</td>
</tr>
<tr>
<td>• The company has the customers’ best interests at heart</td>
</tr>
<tr>
<td>• The employees of the company understand the specific needs of their customers</td>
</tr>
</tbody>
</table>

Study Aims and Research Question

Any new service model implementation should usually be well planned to improve user satisfaction; however, home-based teleconsultation at stroke prevention clinics was implemented without the usual planning. The rapid implementation could affect patients’ experiences. As a result, it is vital to evaluate patient satisfaction with the quality of the teleconsultation service they received during the COVID-19 pandemic by assessing satisfaction with various service dimensions to identify aspects of service quality that patients are and are not satisfied with. This study aimed to explore the factors affecting patient satisfaction. The research question was as follows: “What are the patient-identified factors influencing patients’ satisfaction with service quality in stroke prevention clinics’ home-based teleconsultation service?”

Methods

Participants and Procedure

We conducted a web-based or telephone survey of patients who had at least one home-based teleconsultation, either the initial or follow-up visit, at the stroke prevention clinics. The study sites were 2 stroke prevention clinics at 2 tertiary hospitals in Ontario, Canada. The study sample consisted of individuals who received at least one home-based teleconsultation at a stroke prevention clinic between January 1, 2021, and November 30, 2021. A convenience sampling technique was used as we invited patients who lived in their homes and self-participated in the stroke prevention clinic home-based teleconsultation service during the COVID-19 pandemic. Our exclusion criteria included (1) patients who lived in a long-term care home or group home and (2) patients with dementia who could not participate in home-based teleconsultations.
To minimize volunteer bias and increase the response rate, we applied various data collection techniques to capture participants with and without internet access. A web survey was used for patients with email addresses, whereas a telephone survey was used for patients without email addresses. The questionnaire was administered between May 17, 2021, and December 10, 2021.

We developed a telephone script for recruitment to explain the research project in lay terms. In total, 2 modified-duty nurses from one site and neurologists from another site who were not part of the research team obtained permission from patients to be contacted by the research team. The list of email addresses or telephone numbers of patients who gave permission was shared with the research team. Participants who chose a web-based survey received an email with a brief cover letter explaining the study’s purpose and their rights as study participants. Informed consent was via the web before accessing the questionnaire, and they were asked to click a box indicating that they agreed to complete the survey (Multimedia Appendix 1). Participants who chose a telephone survey received a mail-in cover letter, a consent form, and a copy of the survey (Multimedia Appendix 2).

**Ethical Considerations**

This study was reviewed by the research ethics boards of Southlake Regional Health Center and Mackenzie Health and was considered a continuous quality improvement project; thus, a full research ethics review was not required. This study was also reviewed by the University of Waterloo Office of Research Ethics (ORE 42686) and received ethical clearance.

**Measures**

Our study used SERVPERF to design a Likert-scale survey to assess patient satisfaction with home-based teleconsultation service quality in stroke prevention clinics. We acknowledge that patient satisfaction is subjective, with many determinants that may not be related to the SERVPERF model. The literature has indicated that patient satisfaction can be influenced by patient knowledge and expectations; therefore, other factors such as demographics (eg, age, gender, and education), clinical factors (eg, comorbidities, diagnosis, and number of visits), and experiences with teleconsultation can influence patient expectations [20,21]. We also included these factors in our survey. By considering other factors and applying direct and indirect measurements, we attempted to explore patient satisfaction using a holistic approach. The 18-item questionnaire used in our study was developed by referencing the telehealth service quality questionnaire developed by Yin et al [30] (see Multimedia Appendix 3 for a description).

The survey consisted of three components: (1) demographic, clinical, and telemedicine questions; (2) an 18-item Likert scale–based questionnaire measured on a 5-point scale, with 1 for strongly disagree and 5 for strongly agree; and (3) 6 open-ended follow-up questions (Multimedia Appendix 4). The demographic, clinical, and telemedicine independent variables were selected based on previous evidence from a literature review and clinical significance from a practice point of view [31]. We conducted a pilot study in March 2021 with 10 participants who had home-based teleconsultation from October 2020 to December 2020 and asked 6 additional questions about the survey content (Multimedia Appendix 5). Overall, patients were satisfied with the language and content of the survey, indicated in their feedback.

**Data Collection**

The web survey was conducted through a secured, password-protected REDCap (Research Electronic Data Capture; Vanderbilt University) website hosted at the University of Waterloo that supports research data collection [32]. Skype for Business (Skype Technologies) from the University of Waterloo, with recording and transcription functions, was set up for the research assistant to conduct the telephone survey.

**Statistical Analysis**

We applied quantitative and qualitative analysis to understand patient satisfaction. A binary outcome variable was defined as (1) a low-satisfaction group if the participants chose very unsatisfied, dissatisfied, and neither satisfied nor satisfied with the overall home-based teleconsultation service quality; and (2) a high-satisfaction group if the participants chose satisfied and very satisfied. We used SPSS for Windows (version 28.0.1; IBM Corp) for statistical analysis [33]. The Likert-scale questions were converted to numerical values. Using the item means, we generated each SERVPERF dimension score and an overall questionnaire score for each respondent’s survey. There were 10 demographic, 7 clinical, and 6 technical-related independent variables (see Multimedia Appendix 6 for the definitions). Chi-square tests were used to identify the statistical significance between the categorical independent and binary outcome variables. The point biserial correlation was calculated to identify the correlation between a continuous independent variable and the binary outcome variable. To test the internal reliability of our instrument, we calculated the Cronbach α. As we had a large number of independent variables under consideration, a forward selection model was most suitable [34]. We used statistically significant variables correlating to the binary outcome variable in the stepwise binary logistic regression model, with \( p < .05 \) considered significant.

We used NVivo (QSR International), a software developed to organize and support the analysis of qualitative data. GM coded the entire data set, and ST independently coded 10 random samples. The results were compared and reached an initial 87.1% agreement. Discrepancies were discussed, and conflicts were resolved after further clarification of the definition of the codes. We applied direct content analysis to understand the service quality of the teleconsultation under study [35]. We used the 5 service dimensions and their operational definitions as the initial coding categories [36]. Next, we read each transcript and identified and categorized all the text that appeared to represent the operational definition of the code [35]. Text not categorized using the initial coding scheme would be considered for a new code. We summarized the categories of the entire data set and then divided them into low- and high-satisfaction groups to explore positive and negative patient perceptions. We compared the differences between the 2 groups that could explain the quantitative analysis results [37].
Results

Participant Characteristics

The response rate was 35.9% (104/290) for the web-based survey and 44% (24/54) for the telephone survey. A total of 110 (n=86, 78.2% web and n=24, 21.8% telephone) surveys were included for quantitative analysis, and 97 (n=74, 76% web and n=23, 24% telephone) surveys were included for direct content analysis. Figures 1 and 2 show a flowchart summarizing the subsequent exclusion of cases from the original number participants to arrive at the final analysis. A total of 97.3% (107/110) of the participants used telephone consultations. The percentages of missing values (1% to 8%) for each Likert-scale question were insignificant (<20%); therefore, the mean of each item was used to replace the missing data [38].

Figure 1. Sample flowchart for the web-based survey.
The descriptive statistics of the demographic, clinical, and telemedicine variables are presented in Tables 1-3. Briefly, most of the participants (99/110, 90%) were aged ≥55 years, were retired (80/109, 73.4%) and married (76/110, 69.1%), lived with others (85/109, 78%), and lived within 20 km of where the stroke prevention clinic was located (77/110, 70%). Only a few participants (7/109, 6.4%) had an educational level lower than high school, and most (67/109, 61.5%) had a postsecondary education (Table 1). Regarding clinical factors, most participants had a stroke diagnosis (71/110, 64.5%) and self-identified as having only one stroke risk factor (74/110, 67.3%). Most patients were new (101/110, 91.8%) to the stroke prevention clinics (Table 2). Many participants had relatively less experience with health technology. Although most of them owned digital equipment for teleconsultation (79/109, 72.5%), many of them had never used patient portals (94/108, 87%) and telemedicine (92/108, 85.2%) before the COVID-19 pandemic (Table 3).
Table 1. Demographic characteristics of the patients included in the study (N=110).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total(^a)</th>
<th>Low satisfaction (n=26)</th>
<th>High satisfaction (n=84)</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group (y), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.04(^b)</td>
</tr>
<tr>
<td>&lt;55</td>
<td>11 (10)</td>
<td>2 (7.7)</td>
<td>9 (10.7)</td>
<td></td>
</tr>
<tr>
<td>55-64</td>
<td>18 (16.4)</td>
<td>3 (11.5)</td>
<td>15 (17.8)</td>
<td></td>
</tr>
<tr>
<td>65-74</td>
<td>25 (25.5)</td>
<td>4 (15.4)</td>
<td>24 (28.6)</td>
<td></td>
</tr>
<tr>
<td>75-84</td>
<td>40 (36.4)</td>
<td>16 (61.5)</td>
<td>24 (28.6)</td>
<td></td>
</tr>
<tr>
<td>≥85</td>
<td>13 (11.8)</td>
<td>1 (3.8)</td>
<td>12 (14.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.52</td>
</tr>
<tr>
<td>Female</td>
<td>49 (44.5)</td>
<td>13 (50)</td>
<td>36 (42.9)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>61 (55.5)</td>
<td>13 (50)</td>
<td>48 (57.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Distance (km), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.31</td>
</tr>
<tr>
<td>1-20, n (%)</td>
<td>77 (70)</td>
<td>21 (80.8)</td>
<td>56 (66.7)</td>
<td></td>
</tr>
<tr>
<td>21-40, n (%)</td>
<td>21 (19.1)</td>
<td>4 (15.4)</td>
<td>17 (20.2)</td>
<td></td>
</tr>
<tr>
<td>&gt;40, n (%)</td>
<td>12 (10.9)</td>
<td>1 (3.8)</td>
<td>10 (13.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Education, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.52</td>
</tr>
<tr>
<td>Grade 8 or lower</td>
<td>7 (6.4)</td>
<td>0 (0)</td>
<td>7 (8.3)</td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>35 (32.1)</td>
<td>10 (40)</td>
<td>25 (29.8)</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>32 (29.4)</td>
<td>6 (24)</td>
<td>26 (31)</td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>24 (21.8)</td>
<td>6 (24)</td>
<td>18 (21.4)</td>
<td></td>
</tr>
<tr>
<td>Graduate school</td>
<td>11 (10)</td>
<td>3 (12)</td>
<td>8 (9.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Employment, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.68</td>
</tr>
<tr>
<td>Retired</td>
<td>80 (73.4)</td>
<td>20 (76.9)</td>
<td>60 (72.3)</td>
<td></td>
</tr>
<tr>
<td>Working</td>
<td>17 (15.6)</td>
<td>2 (7.7)</td>
<td>15 (18.1)</td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>4 (3.7)</td>
<td>1 (3.8)</td>
<td>3 (3.6)</td>
<td></td>
</tr>
<tr>
<td>Self-employed</td>
<td>3 (2.8)</td>
<td>1 (3.8)</td>
<td>2 (2.4)</td>
<td></td>
</tr>
<tr>
<td>On disability</td>
<td>5 (4.6)</td>
<td>2 (7.7)</td>
<td>3 (3.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Marital status, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.42</td>
</tr>
<tr>
<td>Married</td>
<td>76 (69.1)</td>
<td>18 (69.2)</td>
<td>58 (69.1)</td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>11 (10)</td>
<td>4 (15.4)</td>
<td>7 (8.3)</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>17 (15.5)</td>
<td>4 (15.4)</td>
<td>13 (15.5)</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>6 (5.5)</td>
<td>0 (0)</td>
<td>6 (7.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Living arrangement, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.69</td>
</tr>
<tr>
<td>With others</td>
<td>85 (78)</td>
<td>21 (80.8)</td>
<td>64 (77.1)</td>
<td></td>
</tr>
<tr>
<td>Alone</td>
<td>24 (22)</td>
<td>5 (19.2)</td>
<td>19 (22.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Transportation, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.24</td>
</tr>
<tr>
<td>Drives</td>
<td>67 (65)</td>
<td>12 (54.5)</td>
<td>55 (67.9)</td>
<td></td>
</tr>
<tr>
<td>Relies on others</td>
<td>36 (35)</td>
<td>10 (45.5)</td>
<td>26 (32.1)</td>
<td></td>
</tr>
<tr>
<td>Use of a cane or walker (yes), n (%)</td>
<td>30 (27.5)</td>
<td>8 (30.8)</td>
<td>22 (26.2)</td>
<td>.57</td>
</tr>
<tr>
<td>Language barrier (yes), n (%)</td>
<td>10 (9.1)</td>
<td>6 (23.1)</td>
<td>4 (4.8)</td>
<td>.005(^c)</td>
</tr>
<tr>
<td>Hearing impaired (yes), n (%)</td>
<td>28 (25.5)</td>
<td>9 (34.6)</td>
<td>19 (22.6)</td>
<td>.22</td>
</tr>
<tr>
<td>Affecting phone conversations (yes)</td>
<td>14 (50)</td>
<td>6 (66.7)</td>
<td>8 (42.1)</td>
<td></td>
</tr>
<tr>
<td>Vision impaired (yes), n (%)</td>
<td>33 (30)</td>
<td>8 (30.8)</td>
<td>25 (29.8)</td>
<td>.92</td>
</tr>
<tr>
<td>Variable</td>
<td>Total</td>
<td>Low satisfaction (n=26)</td>
<td>High satisfaction (n=84)</td>
<td>P value</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>------------------------</td>
<td>-------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Affecting the use of a screen (yes)</td>
<td>16 (50)</td>
<td>5 (62.5)</td>
<td>11 (45.8)</td>
<td>.41</td>
</tr>
<tr>
<td>The survey was consented to by the SDM&lt;sup&gt;d&lt;/sup&gt; (yes), n (%)</td>
<td>53 (49.1)</td>
<td>18 (69.2)</td>
<td>35 (41.7)</td>
<td>.01&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Web survey (yes), n (%)</td>
<td>86 (78.2)</td>
<td>24 (92.3)</td>
<td>62 (73.8)</td>
<td>.046&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Note that the percentages are based on denominators that vary from the overall sample size of 110 because of missing data.

<sup>b</sup>P<.05.

<sup>c</sup>P<.01.

<sup>d</sup>SDM: substitute decision maker. Indicates that the survey was consented to and answered with the help of an SDM.

### Table 2. Clinical characteristics of the patients included in the study (N=110).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total, n (%)</th>
<th>Low satisfaction (n=26), n (%)</th>
<th>High satisfaction (n=84), n (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Had stroke diagnosis</td>
<td>71 (64.5)</td>
<td>17 (65.4)</td>
<td>54 (64.3)</td>
<td>.92</td>
</tr>
<tr>
<td>Had stroke residual deficits</td>
<td>27 (38)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>9 (52.9)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>18 (33.3)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>.15</td>
</tr>
<tr>
<td>New referral</td>
<td>101 (91.8)</td>
<td>25 (96.2)</td>
<td>76 (90.5)</td>
<td>.36</td>
</tr>
</tbody>
</table>

| Number of stroke risk factors | | | | |
| 0 | 14 (12.7) | 3 (11.5) | 11 (13.1) | |
| 1 | 74 (67.3) | 21 (80.8) | 53 (63.1) | |
| 2-4 | 19 (17.3) | 1 (3.8) | 18 (21.4) | |
| 5-6 | 3 (2.7) | 1 (3.8) | 2 (2.4) | |

<sup>a</sup>N=71.

<sup>b</sup>N=17.

<sup>c</sup>N=54.

### Table 3. Telemedicine-related characteristics of the patients included in the study (N=110).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total, n (%)</th>
<th>Low satisfaction (n=26), n (%)</th>
<th>High satisfaction (n=84), n (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used telephone</td>
<td>107 (98.2)</td>
<td>26 (100)</td>
<td>81 (97.6)</td>
<td>.42</td>
</tr>
</tbody>
</table>

| Number of stroke prevention clinic home-based teleconsultations | | | | |
| Once only | 52 (47.3) | 13 (50) | 39 (46.4) | .74 |
| 2-4 times | 50 (45.5) | 12 (46.2) | 38 (45.2) | |
| ≥5 times | 8 (7.3) | 1 (3.8) | 7 (8.3) | |

| Patient portal use before the COVID-19 pandemic | | | | |
| Never | 94 (85.5) | 23 (92) | 71 (85.5) | .71 |
| 1-2 times | 10 (9.1) | 2 (8) | 8 (9.6) | |
| 3-5 times | 3 (2.7) | 0 (0) | 3 (3.6) | |
| >5 times | 1 (0.9) | 0 (0) | 1 (1.2) | |

| Telemedicine use before the COVID-19 pandemic | | | | |
| Never | 92 (83.6) | 21 (80.8) | 71 (86.6) | .35 |
| 1-2 times | 7 (6.4) | 2 (7.7) | 5 (6.1) | |
| 3-5 times | 6 (5.5) | 1 (3.8) | 5 (6.1) | |
| >5 times | 3 (2.7) | 2 (7.7) | 1 (1.2) | |
| Previsit contact by the stroke prevention clinic (no) | 94 (85.5) | 24 (92.3) | 70 (85.4) | .36 |

| Owned digital equipment at home (yes) | 79 (71.8) | 19 (73.1) | 60 (72.3) | .94 |

<sup>a</sup>Note that the percentages are based on denominators that vary from the overall sample size of 110 owing to missing data.
Findings From the Quantitative Analysis

Overall, the instrument was reliable as the Cronbach α reliability analysis of the SERVPERF questionnaire was .894 (Multimedia Appendix 7), which indicated an excellent level of reliability of the instrument [39]. The adjusted $R^2$ value was 0.76, indicating that the 5 SERVPERF dimensions could explain 76% of the variation in the global satisfaction score. The mean global satisfaction score was 2.5 (SD 0.65) for the low-satisfaction group and 4.40 (SD 0.49) for the high-satisfaction group. To examine the explanatory variables that were significantly associated with the binary outcome variable, consent from the substitute decision maker (SDM), language barrier, age group, survey method, and 5 service dimensions were entered into the final forward stepwise logistic regression model. The adjusted $R^2$ indicated that 69% of the variance could be explained in the final model. Table 4 illustrates that consent from the SDM (adjusted odds ratio [AOR] 6.59, 95% CI 1.45-29.93; $P=.01$) was positively associated ($\beta=1.89$) and the responsiveness (AOR 0.03, 95% CI 0.006-0.188; $P<.001$) and empathy (AOR 0.12, 95% CI 0.02-0.80; $P=.03$) dimensions were negatively associated ($\beta=-3.37$ for responsiveness; $\beta=-2.15$ for empathy) with dissatisfaction with the home-based teleconsultation service quality (Table 4). The odds of dissatisfaction for participants who consented to the survey through their SDM were 6.59 (95% CI 1.45-29.93) compared with those who consented themselves. Every one-unit increase in the responsiveness dimension score decreased the odds of dissatisfaction by 0.03 (95% CI 0.006-0.19) when other variables were held constant. Every one-unit increase in the empathy dimension score decreased the odds of dissatisfaction by 0.12 (95% CI 0.02-0.8) when other variables were held constant.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta$ (SE)</th>
<th>$P$ value</th>
<th>AOR$^a$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consent from SDM$^b$</td>
<td>1.89 (0.772)</td>
<td>.01$c$</td>
<td>6.59 (1.45-29.93)</td>
</tr>
<tr>
<td>Responsiveness$^d$</td>
<td>-3.37 (0.867)</td>
<td>&lt;.001$^e$</td>
<td>0.03 (0.006-0.19)</td>
</tr>
<tr>
<td>Empathy$^f$</td>
<td>-2.15 (0.986)</td>
<td>.03$^c$</td>
<td>0.12 (0.02-0.80)</td>
</tr>
</tbody>
</table>

$^a$AOR: adjusted odds ratio.
$^b$SDM: substitute decision maker. Consent from SDM indicates that the survey was answered with the help of an SDM.
$^cP<.05$.
$^d$Responsiveness is a service performance model dimension regarding the willingness to help customers and provide a prompt service.
$^eP<.001$.
$^f$Empathy is a service performance model dimension regarding providing individual care and attention to customers.

The significant characteristics of the participants who had their SDM sign the consent form and help them answer the survey are listed in Table 5. The participants whose SDM provided consent to help them answer the survey were more likely to answer a web-based survey ($\chi^2=15.6; P<.001$) and have a language barrier ($\chi^2=15.6; P<.001$), hearing impairment ($\chi^2=3.9; P=.048$), or hearing that affected telephone conversations ($\chi^2=5.6; P<.02$) and were less likely to drive ($\chi^2=7.0; P=.04$). Tangibles was the only statistically significant SERVPERF dimension ($P<.001$). The participants who consented through their SDM had a shorter travel distance, were older, and were more likely to have residual stroke symptoms.
### Table 5. Comparison of variables between surveys for which consent was obtained from the substitute decision maker (SDM) and surveys for which consent was obtained from the patient themselves (difference of >15%).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Consent from SDM(^a) (n=53)</th>
<th>Consent from patient (n=57)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web-based survey, n (%)</td>
<td>50 (94)</td>
<td>36 (63)</td>
<td>&lt;.001(^b)</td>
</tr>
<tr>
<td>Age (y), mean (SD)</td>
<td>73.75 (13.085)</td>
<td>70.60 (10.61)</td>
<td>.17</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>34 (64)</td>
<td>27 (47)</td>
<td>.08</td>
</tr>
<tr>
<td>Distance (km), mean (SD)</td>
<td>15.897 (16.525)</td>
<td>19.435 (16.48)</td>
<td>.26</td>
</tr>
<tr>
<td>Driving, n (%)</td>
<td>27 (55)(^c)</td>
<td>40 (74)(^d)</td>
<td>.04(^e)</td>
</tr>
<tr>
<td>Language barrier, n (%)</td>
<td>10 (19)</td>
<td>0 (0)</td>
<td>&lt;.001(^b)</td>
</tr>
<tr>
<td>Hearing impaired, n (%)</td>
<td>18 (34)</td>
<td>10 (18)</td>
<td>.048</td>
</tr>
<tr>
<td>Hearing impairment affecting phone conversations, n (%)</td>
<td>12 (67)(^f)</td>
<td>2 (20)(^g)</td>
<td>.02(^e)</td>
</tr>
<tr>
<td>Residual stroke symptoms, n (%)</td>
<td>17 (49)(^h)</td>
<td>10 (28)(^i)</td>
<td>.07</td>
</tr>
<tr>
<td>Tangibles, mean (SD)</td>
<td>3.21 (0.83)</td>
<td>3.67 (0.61)</td>
<td>&lt;.001(^b)</td>
</tr>
<tr>
<td>Reliability, mean (SD)</td>
<td>3.87 (0.82)</td>
<td>3.94 (0.56)</td>
<td>.68</td>
</tr>
<tr>
<td>Responsiveness, mean (SD)</td>
<td>3.76 (0.94)</td>
<td>3.76 (0.59)</td>
<td>.98</td>
</tr>
<tr>
<td>Assurance, mean (SD)</td>
<td>3.99 (0.78)</td>
<td>4.04 (0.53)</td>
<td>.68</td>
</tr>
<tr>
<td>Empathy, mean (SD)</td>
<td>3.75 (0.71)</td>
<td>3.96 (0.56)</td>
<td>.34</td>
</tr>
</tbody>
</table>

\(a\)Consent from SDM indicates that the survey was answered with the help of an SDM.

\(^b\)P < .001.

\(^c\)N=49.

\(^d\)N=54.

\(^e\)P < .05.

\(^f\)N=18.

\(^g\)N=10.

\(^h\)N=35.

\(^i\)N=36.

### Findings From the Content Analysis

#### Overview

A total of 88.2% (97/110) of patients completed the open-ended questions in the survey, with 25% (24/97) in the low-satisfaction group and 75% (73/97) in the high-satisfaction group. Multimedia Appendices 8 and 9 list the dimensions and subcategories of positive and negative comments among patients with low and high satisfaction scores, respectively. Interestingly, the low- and high-satisfaction groups shared the same dissatisfied service dimensions (assurance, reliability, and empathy) and subcategories. Overall, missing clinical components, inadequate communication, administrative issues, and absence of personal connection were the significant concerns that affected patients’ perceived home-based teleconsultation quality at the stroke prevention clinics.

In contrast, the most satisfying service dimensions were assurance, empathy, and responsiveness among the high-satisfaction group. Overall, a competent clinician with effective communication skills and great empathy for patients is crucial for patient-perceived high-quality care in home-based teleconsultation. In addition, convenience, appropriateness to the patient’s situation, and timely consultation were important for high satisfaction with home-based teleconsultation.

### Future Use of Home-Based Teleconsultation

Assessment of patient preference for future use of home-based teleconsultation under normal circumstances (after the COVID-19 pandemic) showed that 35% (33/95) of participants preferred not to use home-based teleconsultation. Most participants (18/23, 78%) in the low-satisfaction group preferred not to use home-based teleconsultation. In contrast, 56% (40/72) of participants in the high-satisfaction group were willing to use it, and 24% (17/72) indicated that they might use home-based teleconsultation for specific reasons. However, a minority of participants in the high-satisfaction group (15/72, 21%) still refused to use it in normal circumstances, with primary concerns of communication issues, lack of personal connection, and the belief in the superiority of in-person consultations.

### Discussion

#### Principal Findings

Home-based teleconsultation as a form of telemedicine rapidly expanded in many health sectors in Canada during the
COVID-19 pandemic owing to lockdowns and social distancing restrictions [31]. Since the COVID-19 pandemic, home-based teleconsultation has become essential in outpatient service delivery [40]. By April 2020, 77% of Ontario ambulatory visits were conducted using teleconsultation, a total of 77% of ambulatory visits were conducted using a virtual modality [41]. Nearly all (32/33, 97%) Ontario stroke prevention clinics that responded to a province-wide survey from June 2021 to July 2021 (response rate of 33/41, 80%) reported that they had adopted home-based teleconsultation as a service delivery mode in addition to in-person visits since the COVID-19 pandemic [42]. Patient satisfaction should be the priority in future virtual care development and adoption [43]. To our knowledge, this is the first study to use a service quality theoretical lens to assess patient satisfaction with home-based teleconsultation in outpatient stroke care during the pandemic. Many studies of patient surveys during the COVID-19 pandemic have found that most patients and clinicians reported positive experiences with teleconsultation at outpatient neurology services during the COVID-19 pandemic [31]. However, no study has investigated outpatient stroke prevention services or examined the service quality of home-based teleconsultation during the COVID-19 pandemic. The patient population of stroke prevention clinics and the disease characteristics differ from those of other chronic neurological diseases. For instance, patients at stroke prevention clinics have a unique mix of acuity (such as early identification of large vessel occlusion and cardiac source of embolism) and chronic disease management (eg, hypertension, dyslipidemia, diabetes, and lifestyle management), and most are older adults. Owing to health resource disparity, timely access to outpatient magnetic resonance imaging is not always feasible for minor strokes. When referred to stroke prevention clinics, patients with TIA have transient neurological symptoms, unremarkable brain images, and normal physical examinations. History taking is essential in patients with TIA. The unique patient population and characteristics may pose different challenges in home-based teleconsultation, especially for newly referred patients. Our study found that the participants who were older (mean age 72.12, SD 11.92 years) and mostly newly referred (101/110, 91.8%) and used the telephone modality (107/109, 98.2%) were satisfied with the home-based teleconsultation provided by the stroke prevention clinics during the COVID-19 pandemic. We identified patient-reported factors that affected their satisfaction with the service quality of home-based teleconsultation. Our study filled these research gaps.

Responsiveness was the most statically significant dimension in our quantitative results and is an influential factor for a positive experience. Convenience was the main subtheme of the responsiveness dimension in the high-satisfaction group. The patients in the low-satisfaction group tended to live closer to the stroke prevention clinics, with an average distance of 13.53 (SD 11.57) km, than those in the high-satisfaction group (mean 19.03, SD 17.61 km). Even though they were less likely to drive and had some communication barriers (more likely to have language barriers or hearing impairments), convenience was not a positive factor influencing their satisfaction. Our content analysis supported that convenience, by saving time, travel, and energy, influenced patients’ positive perceptions of the personal benefits of home-based teleconsultation during the pandemic [31]. Our findings were consistent with those of studies conducted before the COVID-19 pandemic [44]. A systematic review of digital experience also found that convenience is one of the motivating factors contributing to a positive digital patient experience [45]. The convenience of home-based teleconsultation is an influential factor in swaying patients’ satisfaction with service quality to the positive side and their preference for teleconsultation [31]. However, convenience is not equivalent to good quality of care. We need to consider the effect of convenience when assessing patients’ preferences and evaluating patient satisfaction with the service quality of home-based teleconsultation.

Second, the empathy dimension was a significant factor in both the quantitative and qualitative analyses. Dissatisfaction feedback in the empathy dimension was found in both the low- and high-satisfaction groups. Our study participants used mostly the telephone modality, where the lack of nonverbal cues may be associated with a profound concern about the lack of personal connection, which is the primary subcategory of the empathy dimension. As indicated by existing studies, the replacement of interpersonal connection and a lack of physical human contact are negatively associated with digital patient experiences [45]. The literature shows that empathy can have powerful effects on positive patient outcomes and satisfaction [46,47]. There is a concern that the digitalization of health care services could primarily lead to a decrease in the expression of empathy [46]. A study on patient satisfaction with tele-OBSTETRIC care found that a desire for personal connection via face-to-face interaction with a clinician was a critical motivation for selecting in-person versus teleconsultation care modalities [48]. Our findings are in line with those of the previous literature. The lack of personal connection in our content analysis could explain the negative relationship between the empathy dimension and dissatisfaction. An interesting finding from our content analysis was that even 98% (81/83) of the participants in the high-satisfaction group used the telephone modality; they expressed overwhelmingly more positive than negative comments (34 vs 12) on the empathy dimension. The clinician’s empathy skills may significantly enhance patient experiences even with a low-technology modality. Empathy skill training for clinicians, primarily through computer-mediated communications, is a key area to study in the future [46].

Third, although the assurance dimension was not found to be statistically significant between the low- and high-satisfaction groups, this was likely due to no real difference between the 2 groups. Our content analysis indicated that the assurance dimension was one of the most important SERVPERF dimensions in both the low- and high-satisfaction groups. The subcategories raised from the content analysis revealed that the patients’ concerns in the assurance dimension were the missing clinical components—especially physical examination—and inadequate communication. The view of the inferior quality of a remote examination among clinicians was dominant in outpatient neurology teleconsultation before the COVID-19 pandemic [49]. This is likely why most teleconsultations were performed for follow-up patients with chronic neurological diseases before the COVID-19 pandemic [50]. Compared with
before the COVID-19 pandemic, home-based teleconsultation has been widely used in both new and follow-up patients at home without the luxury of having a health care professional assisting a teleconsultation since the COVID-19 pandemic [31]. The lack of a physical examination and inadequate communication could impede the clinician’s ability to diagnose and formulate a treatment plan, especially for new patients [31]. In addition, the use of video is more challenging than the use of a telephone because of the rapid adaptation and lack of preparation. According to the Ontario stroke prevention clinic web survey, nearly half of the clinics use only the telephone [42]. Video consultations enable a certain degree of remote examination and may facilitate better communication, whereas telephone-only visits limit clinical assessment and communication. Telephone consultations lack body language and physical prompts, which could negatively affect the communication between the clinician and patient. However, most of the participants in our study used telephone consultations (107/110, 97.3%), and their overall satisfaction was high (3.9/5). Future studies could consider patient satisfaction when using a telephone-only modality for new referrals in this patient population under normal circumstances.

The appropriateness of patient selection is a critical factor in high-quality home-based teleconsultation from patients’ perspectives. Our statistical findings indicated that patients who required help from their SDM to consent and answer the survey were positively associated with dissatisfaction. The SDM may have chosen a web survey, as the participants had difficulty answering a telephone survey because of language barriers, hearing impairments, or communication difficulties from residual stroke symptoms (such as aphasia, apraxia, or mild cognitive impairment; Table 5). Moreover, the patients who needed their SDM to consent and help them answer the survey scored significantly lower in the tangibles dimension, which may indicate that the participants had low comfort levels with technology and technical difficulties even with the telephone modality. The consent from an SDM has many unknown characteristics and requires further exploration in future research.

Similarly, the analysis showed that there were no statistically significant differences in the reliability dimension between the low- and high-satisfaction groups. However, it is important to note that issues primarily related to administration, which fall under the reliability dimension, had a negative impact on both the low- and high-satisfaction groups. This possibility aligns with findings from a scoping review that the lack of proper administrative support has harmed clinicians’ teleconsultation satisfaction [31]. Our findings indicate that it also negatively affects patient satisfaction. Some of the low satisfaction may be due to the abrupt change to teleconsultation because of the COVID-19 pandemic and the lack of clinical and patient preparation. We know that some administrative problems are not unique to home-based teleconsultation, as they occur during in-person visits. Home-based teleconsultation may have increased the workload by keeping pace with the transitioning workflow among telephone, video, and in-person visits, contributing to the maladaptation of home-based teleconsultation [51]. Establishing a new care pathway for home-based teleconsultation may streamline the administrative workflow. The assurance, empathy, and reliability dimensions all had the most negative comments in both the low- and high-satisfaction groups, and the high-satisfaction group had the most positive remarks in the assurance and empathy dimensions in the content analysis, which showed a double-edged effect. This finding might indicate that different key subcategories have buffer effects on improving patient satisfaction with the service quality of home-based teleconsultation. This may reflect the advantage of using both quantitative and qualitative data to provide diverse types of information [37]. By comparing them side by side, the qualitative analysis may provide insights to explain the quantitative findings.

**The Future of Stroke Prevention Clinics’ Service Delivery Mode**

A combination of teleconsultation and in-person visits for outpatient stroke prevention care is the future. Our study showed that 45% (43/95) of participants were willing to use and 18% (17/95) would consider using home-based teleconsultation in future nonpandemic conditions. A study examining patient preference for telehealth for nonemergent health issues after the COVID-19 pandemic concluded that patients were generally willing to use video but preferred in-person visits [52]. A patient-centered service should be delivered by offering the patient a choice [20]. Virtual care provides an opportunity to design a health system that is actually patient-centered [43]. A combination of in-person visits and home-based teleconsultation—a hybrid care model—could best meet patient needs by improving efficiency and capacity without added risk [53,54].

Hybrid care should be sustainable in practice settings to ensure patient care quality, equity, and justice [53]. To avoid increasing the digital divide, a telephone may be favorable instead of video calls for older patients and those with a lower education or income and from racial and ethnic minority groups [55]. However, we should refrain from creating a 2-tiered health care system in which high-income individuals receive video consultations and low-income individuals receive phone consultations. Patients should receive the right care in the right setting, at the right time, and with the right mode; the cost of the service should be reduced; and the best clinical practice guidelines should be followed [56]. Hybrid care could be a balanced approach to achieving a high-performance health care system. Patients can choose the best model by considering flexible options, and clinicians can offer individualized recommendations for optimal care modalities [54].

**Limitations**

Our study has several limitations. First, the cross-sectional survey only provides a snapshot of a phenomenon and cannot determine the temporal relationship between the dependent and independent variables. Second, the participants in this study may not be generalizable to patients of other stroke prevention clinics in Ontario, notably in areas with different health resources such as urban versus very remote rural centers. In addition, the open-ended responses to the web survey were very
brief, limiting our ability to gain a deeper understanding of their experiences. Next, as we surveyed patients who had had a home-based teleconsultation within 6 months, recall bias is possible [57]. Only patients who had a home-based teleconsultation in January 2021 and February 2021 received the survey 4 to 5 months later; the following patients received their survey 2 and a half months after the home-based teleconsultation on average (range 1-3 months). In addition, there is potential nonresponse bias, as web surveys usually have a low response rate [57]. A meta-analysis comparing web survey response rates concluded that the average response rate for web surveys was approximately 11% [58]. Our web survey had a 35.9% (104/290) response rate, and the telephone survey yielded a 44% (24/54) response rate.

Overall, although our study may only reflect part of the concept of satisfaction with service quality because of its complexity, it provided substantial insights into areas for quality improvements from patients’ point of view. The literature suggests that patient satisfaction is a multidimensional concept that still needs to be fully defined. The patient satisfaction scores may reflect the demographic mix and clinical and psychological picture of the patients served by a medical service [59]. Our study attempted to use a theory-guided quantitative and qualitative analysis to reveal the relationship and explanation of such a complex phenomenon. Despite the problems of using patient satisfaction to assess service quality, its measurement provides unique information regarding the care process as seen through the patients’ eyes [59]. Patients still provide the best source of accurate information on the care they receive [60]. In the absence of choice in public-funded health care, our survey gave patients a voice to indicate their preferences [21].

Conclusions

Our findings highlighted 2 crucial service quality dimensions (responsiveness and empathy) that were negatively statistically significantly associated with patient dissatisfaction. Moreover, we identified that a survey consented to by an SDM was positively associated with dissatisfaction. In addition, there were 4 subcategories related to patient dissatisfaction (missing clinical activities, inadequate communication, administrative process issues, and absence of personal connection). We anticipate that appropriate patient selection, consideration of patient preferences, a streamlined home-based teleconsultation administrative workflow, and a competent clinician with communication and empathy skills are essential for achieving high satisfaction with home-based teleconsultation. These factors could be considered when designing home-based teleconsultation services to enhance patient experiences of stroke prevention care.

Acknowledgments

This work was supported by 2 stroke prevention clinics. The authors acknowledge especially Dr Courtney Scott, a neurologist, for helping with patient recruitment and Lorri Reynold, the director of Patient Experience, Professional Practice, and Research, for arranging personnel for patient recruitment. This work was also supported by the University of Waterloo, especially by Gaya Bin Noon, a master’s student at the School of Public Health Sciences of the University of Waterloo, who served as the research assistant for this study.

Data Availability

Data are available upon request from the authors.

Conflicts of Interest

GM is a clinical practitioner in one of the stroke prevention clinics but was not involved in patient contact during the research, including the recruitment and telephone interview processes.

Multimedia Appendix 1
Survey consent sheet for participants.
[DOCX File, 26 KB - cardio_v8i1e51439_app1.docx ]

Multimedia Appendix 2
Cover letter for mail-in package.
[DOCX File, 19 KB - cardio_v8i1e51439_app2.docx ]

Multimedia Appendix 3
The description of 18 items used in the service performance model questionnaire.
[DOCX File, 13 KB - cardio_v8i1e51439_app3.docx ]

Multimedia Appendix 4
[DOCX File, 24 KB - cardio_v8i1e51439_app4.docx ]
Multimedia Appendix 5
Six questions to check the clarity of the survey questions.
[DOCX File, 13 KB - cardio_v8i1e51439_app5.docx]

Multimedia Appendix 6
Definitions of 10 demographic, 7 clinical, and 6 technical-related independent variables.
[DOCX File, 14 KB - cardio_v8i1e51439_app6.docx]

Multimedia Appendix 7
Cronbach α reliability statistics and item-total statistics of the service performance model questionnaire.
[DOCX File, 14 KB - cardio_v8i1e51439_app7.docx]

Multimedia Appendix 8
Positive and negative categories among patients with low global satisfaction.
[DOCX File, 17 KB - cardio_v8i1e51439_app8.docx]

Multimedia Appendix 9
Positive and negative categories among patients with high global satisfaction.
[DOCX File, 18 KB - cardio_v8i1e51439_app9.docx]

References


31. Meng et al. JMI R CARDIO [FREE FULL TEXT] [Medline: 32799379]


44. Fleischhacker CL. Patient satisfaction with telehealth services compared to in-office visits: a systematic literature review. Minnesota State University. 2020. URL: https://cornerstone.lib.mnsu.edu/etds/982/ [accessed 2023-07-21]


58. Meng et al. JMIR CARDIO 2024 | vol. 8 | e51439 | p.83 https://cardio.jmir.org/2024/1/e51439 (page number not for citation purposes)
Factors That Influence Patient Satisfaction With the Service Quality of Home-Based Teleconsultation During the COVID-19 Pandemic: Cross-Sectional Survey Study

JMIR Cardio 2024;8:e51439
URL: https://cardio.jmir.org/2024/1/e51439
doi:10.2196/51439
PMID:38363590

©Guangxia Meng, Carrie McAiney, Ian McKillop, Christopher M Perlman, Shu-Feng Tsao, Helen Chen. Originally published in JMIR Cardio (https://cardio.jmir.org), 16.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cardio, is properly cited. The complete bibliographic information, a link to the original publication on https://cardio.jmir.org, as well as this copyright and license information must be included.